Learning for vision III
Convolutional networks

Karel Zimmermann

http://cmp.felk.cvut.cz/~zimmerk/

Vision for Robotics and Autonomous Systems
https://cyber.felk.cvut.cz/vras/

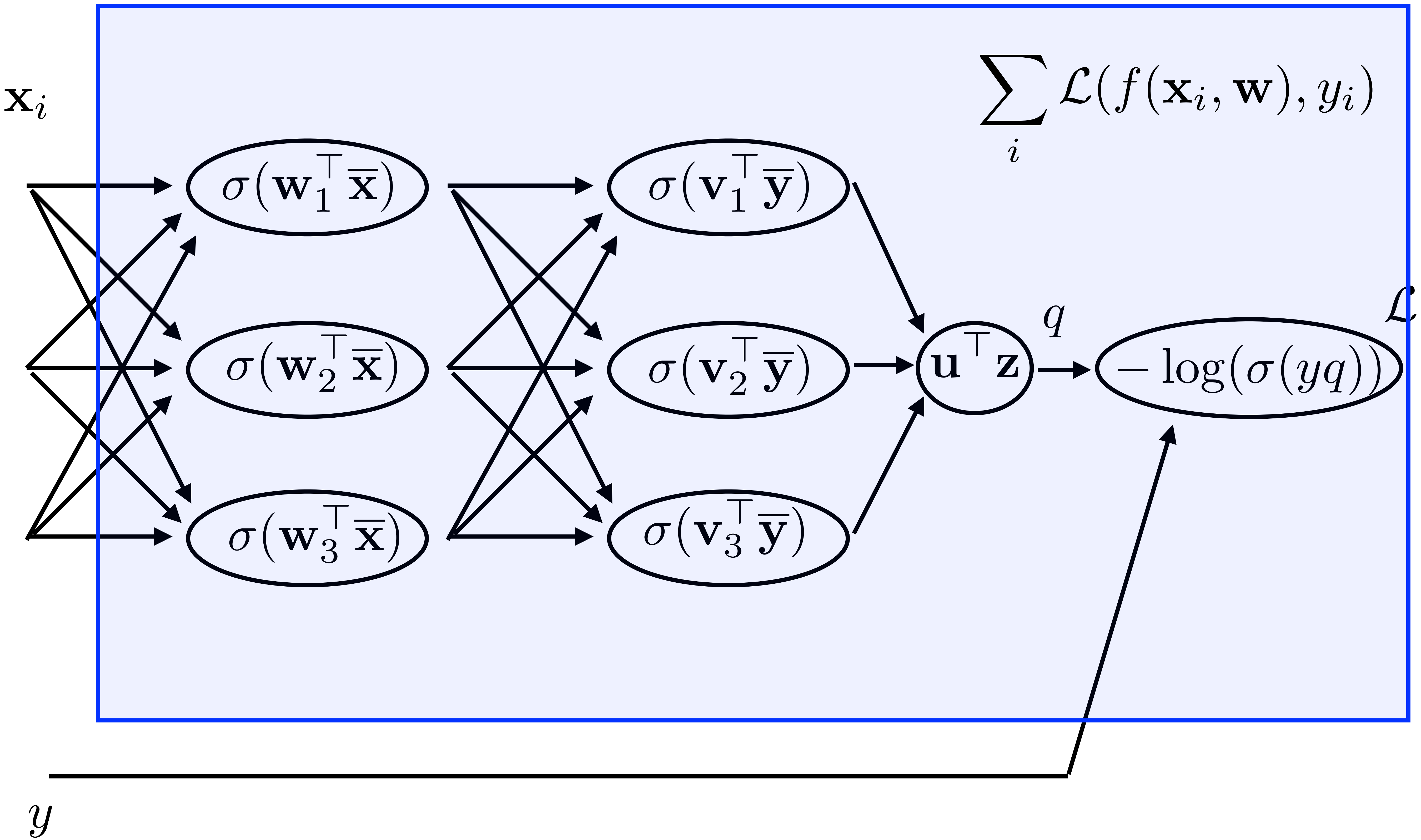
Center for Machine Perception
https://cmp.felk.cvut.cz

Department for Cybernetics
Faculty of Electrical Engineering
Czech Technical University in Prague

Outline

- Fully connected neural network
- Avoid overfitting by search for the NN model suitable for image processing [Hubel and Wiesel 1960].
- Feedforward and Backprop in ConvNets.

Fully connected neural network

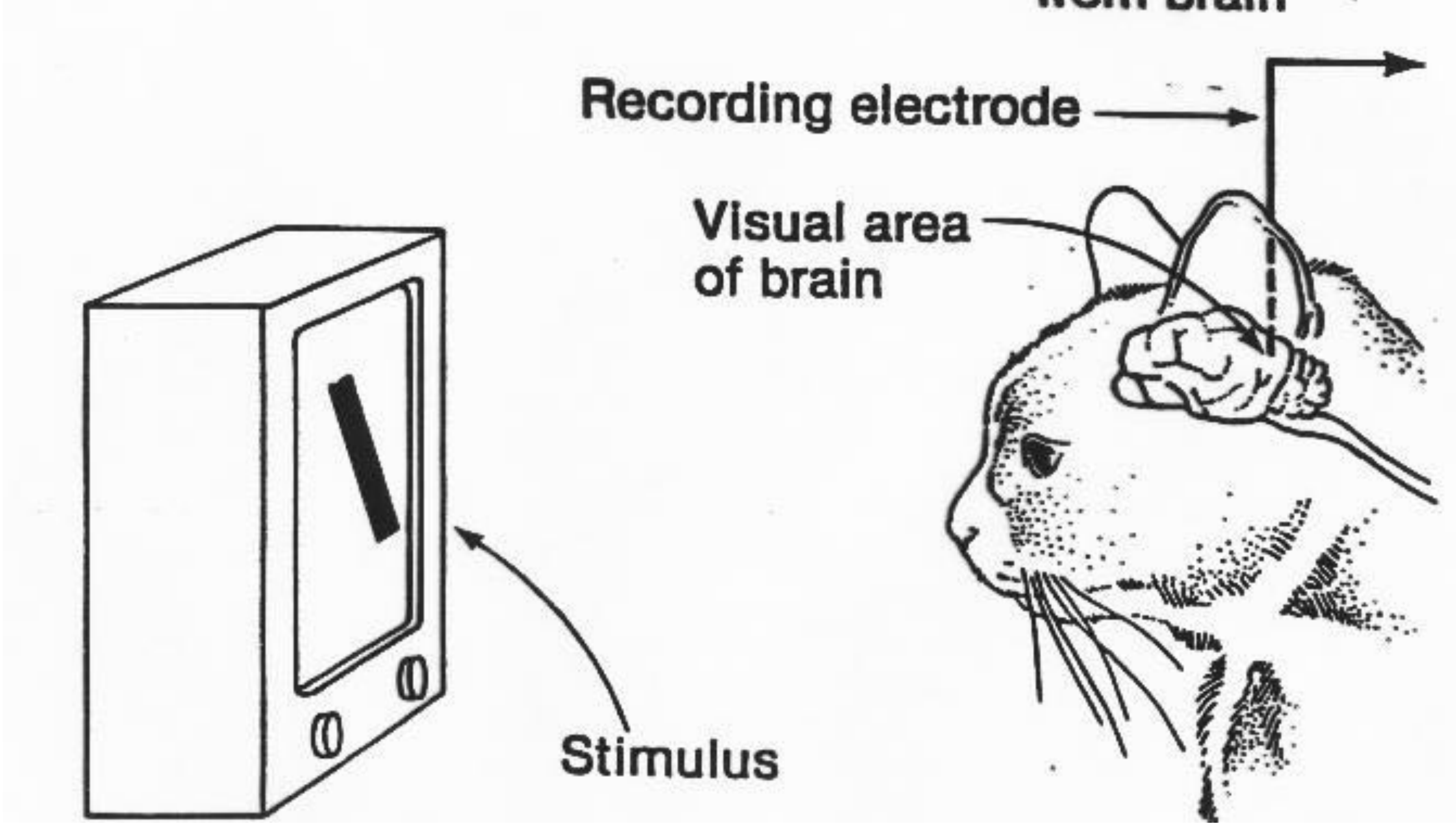


The Tungsten Electrode [Hubel-Science-1957]

http://braintour.harvard.edu/archives/portfolio-items/hubel-and-wiesel

- Device capable to record signal from a single neuron

[Hubel and Wiesel 1959]



- Experiment with anaesthetised paralysed cat

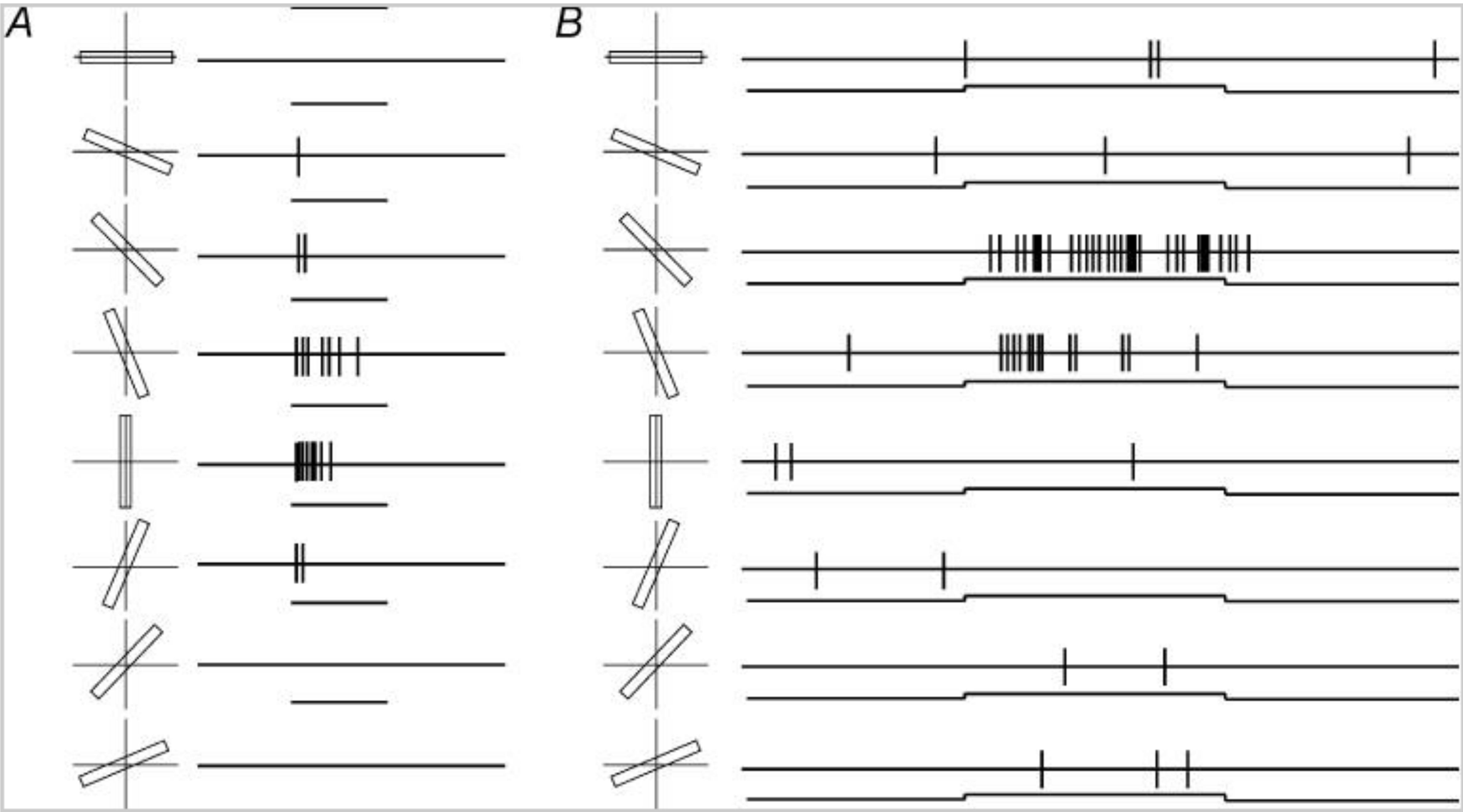
[Hubel and Wiesel 1960]

https://knowingneurons.com/2014/10/29/hubel-and-wiesel-the-neural-basis-of-visual-perception/

[Hubel and Wiesel 1960]

paralysed cat

awake monkey



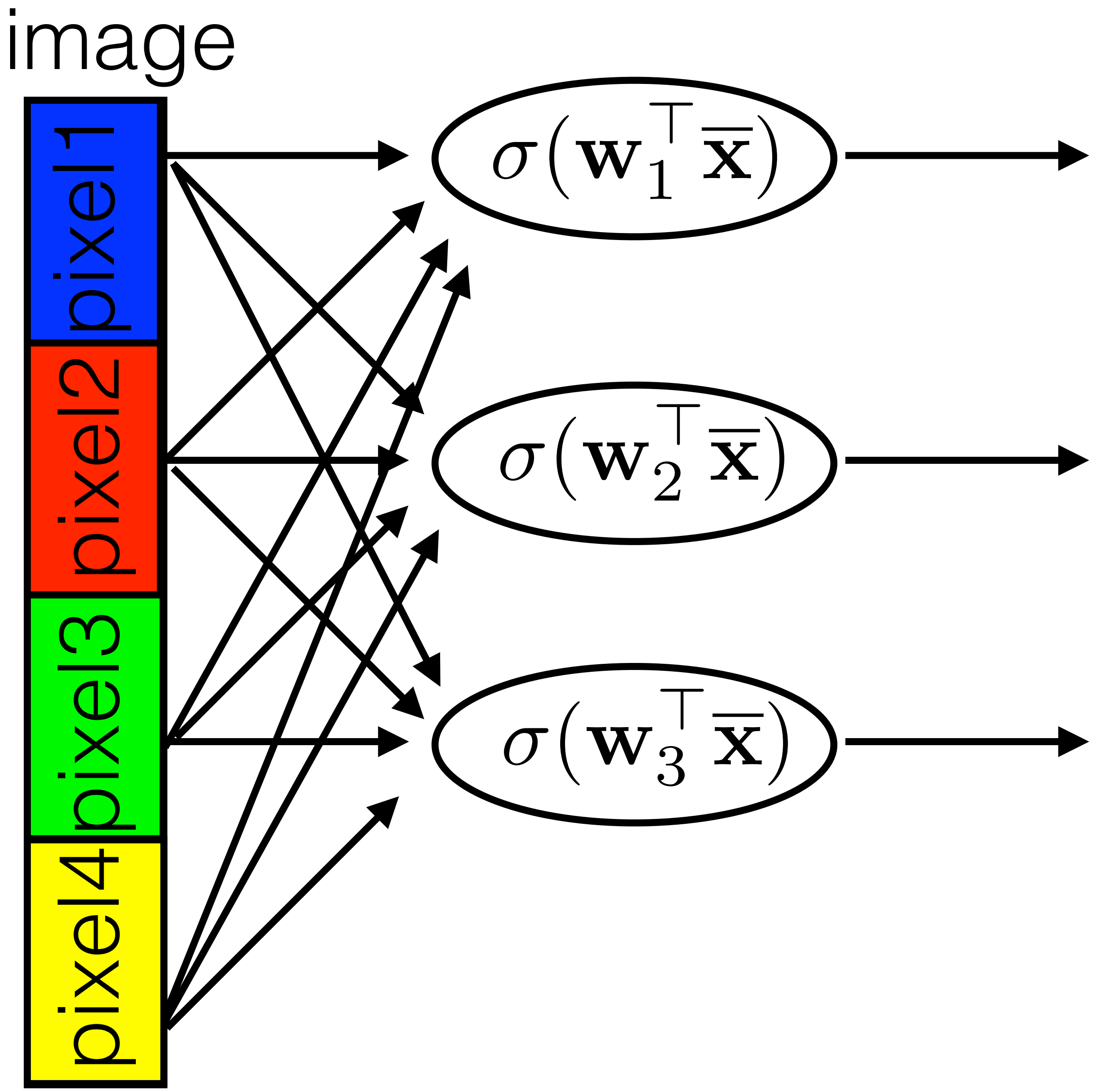
https://knowingneurons.com/2014/10/29/hubel-and-wiesel-the-neural-basis-of-visual-perception/

Hubel and Wiesel experiments in 1950s and 1960s

- Nobel Prize in Physiology and Medicine in 1981
- Dr. Hubel: “There has been a myth that the brain cannot understand itself. It is compared to a man trying to lift himself by his own bootstraps. We feel that is nonsense. The brain can be studied just as the kidney can.”

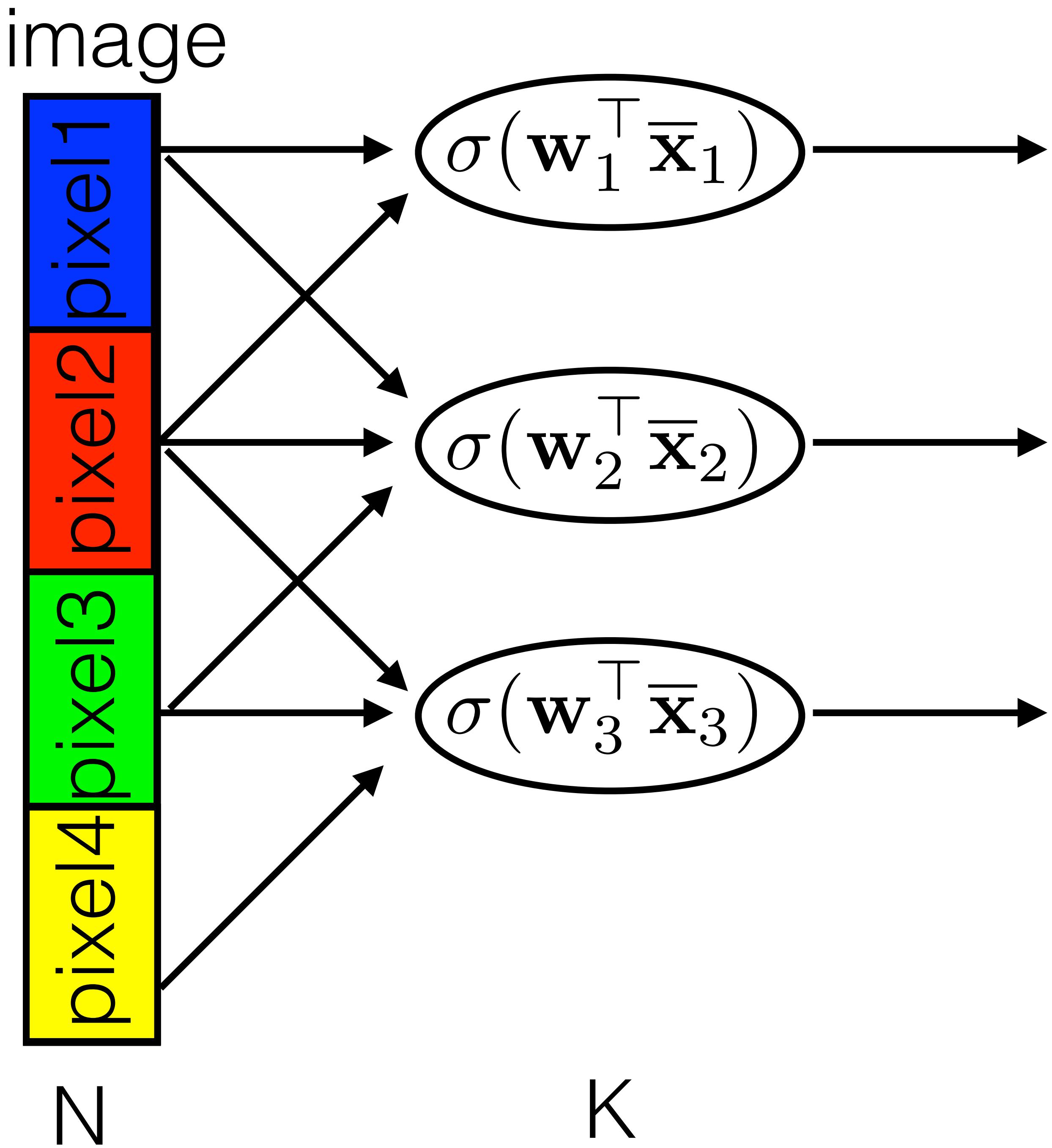
https://knowingneurons.com/2014/10/29/hubel-and-wiesel-the-neural-basis-of-visual-perception/

1. Nearby neurons process information from nearby visual fields (topographical map).



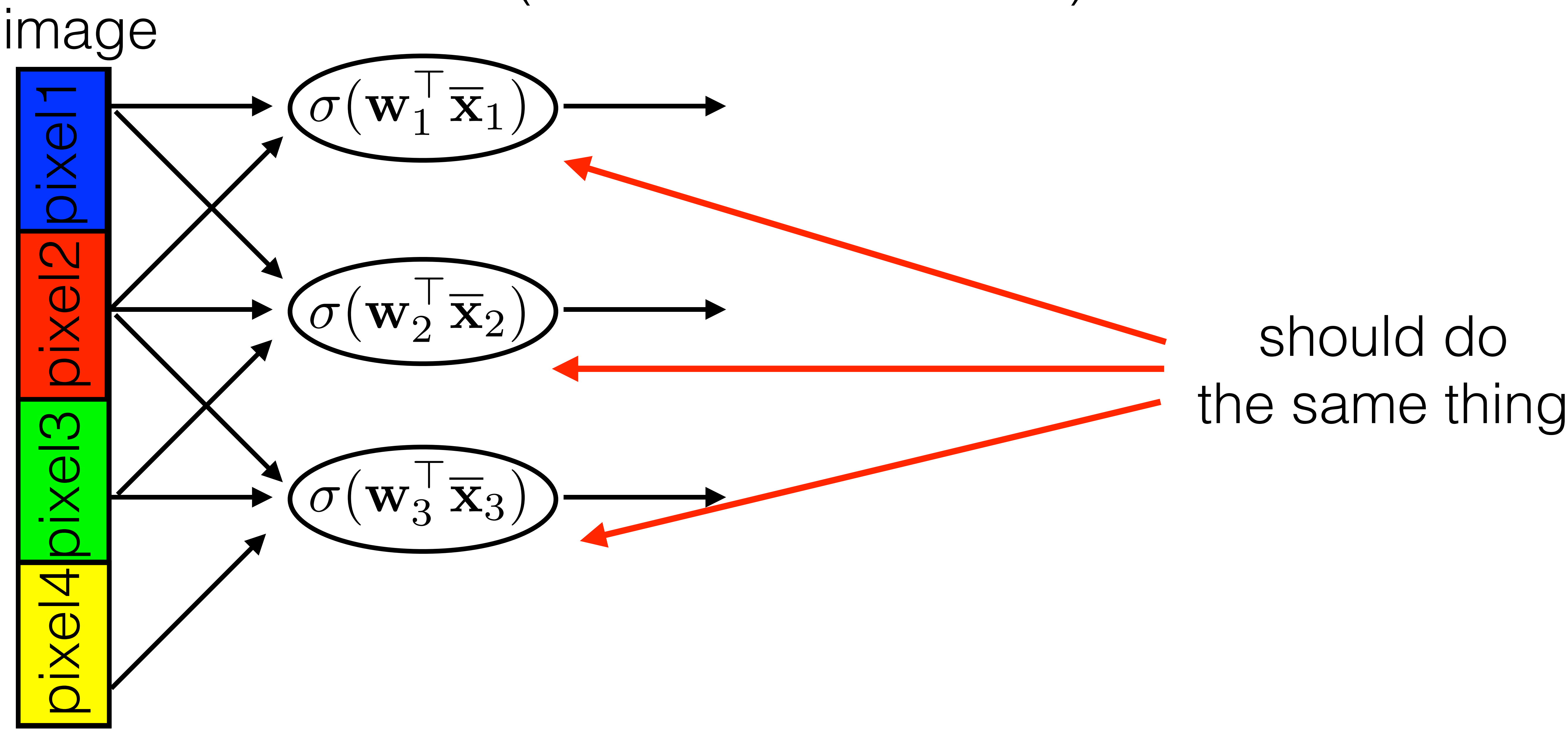
- Processing of visual information in cortex is not fully connected.

1. Nearby neurons process information from nearby visual fields (topographical map).



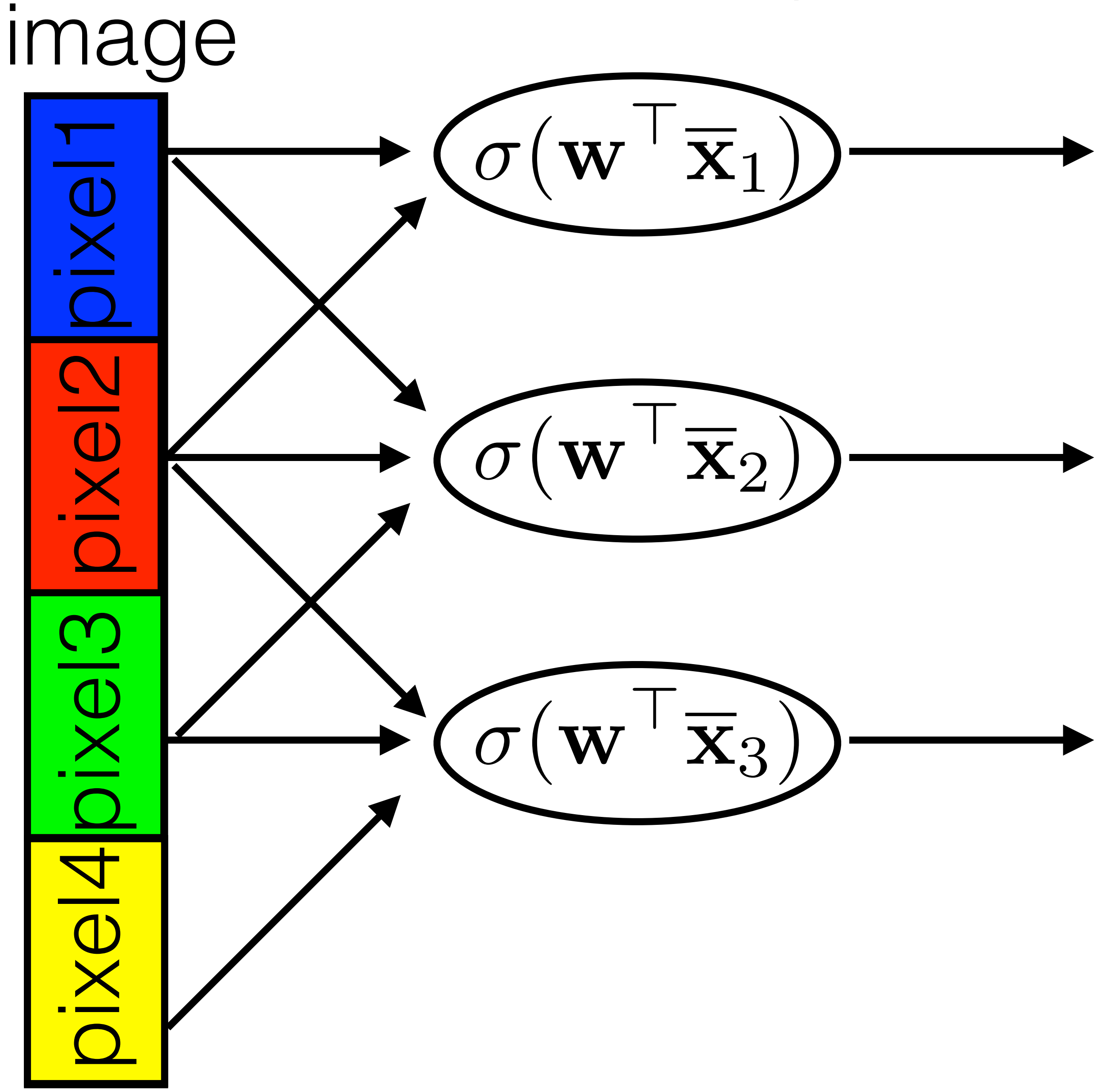
- What is dimensionality reduction for N-pixel image and n-dimensional spatial neighbourhood?

2. Neurons with similar function organized into columns (translation invariance)

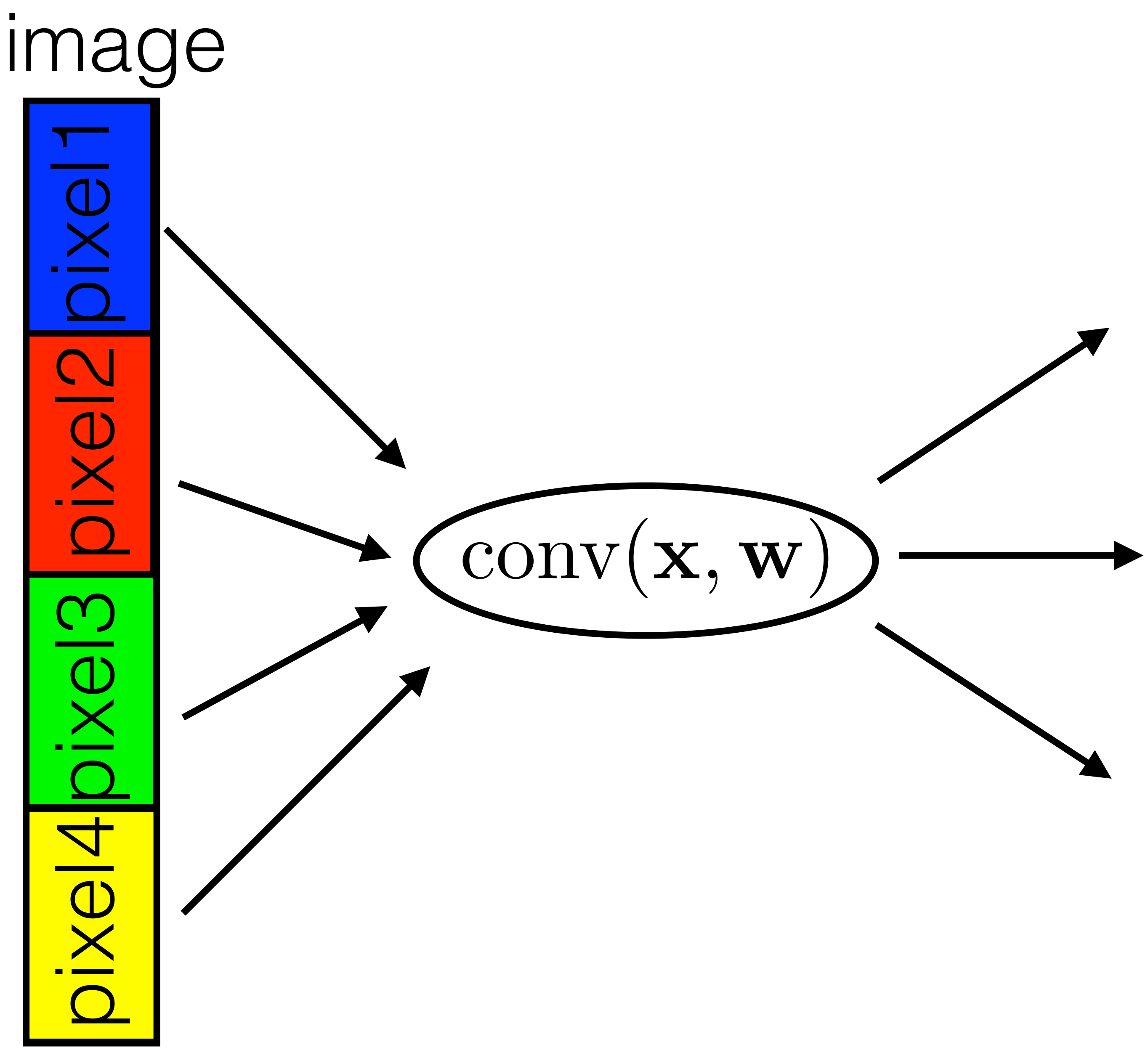


There are neurons which detect an edge on the left and there are different neurons which detect the same edge on the right.

2. Neurons with similar function organized into columns (translation invariance)

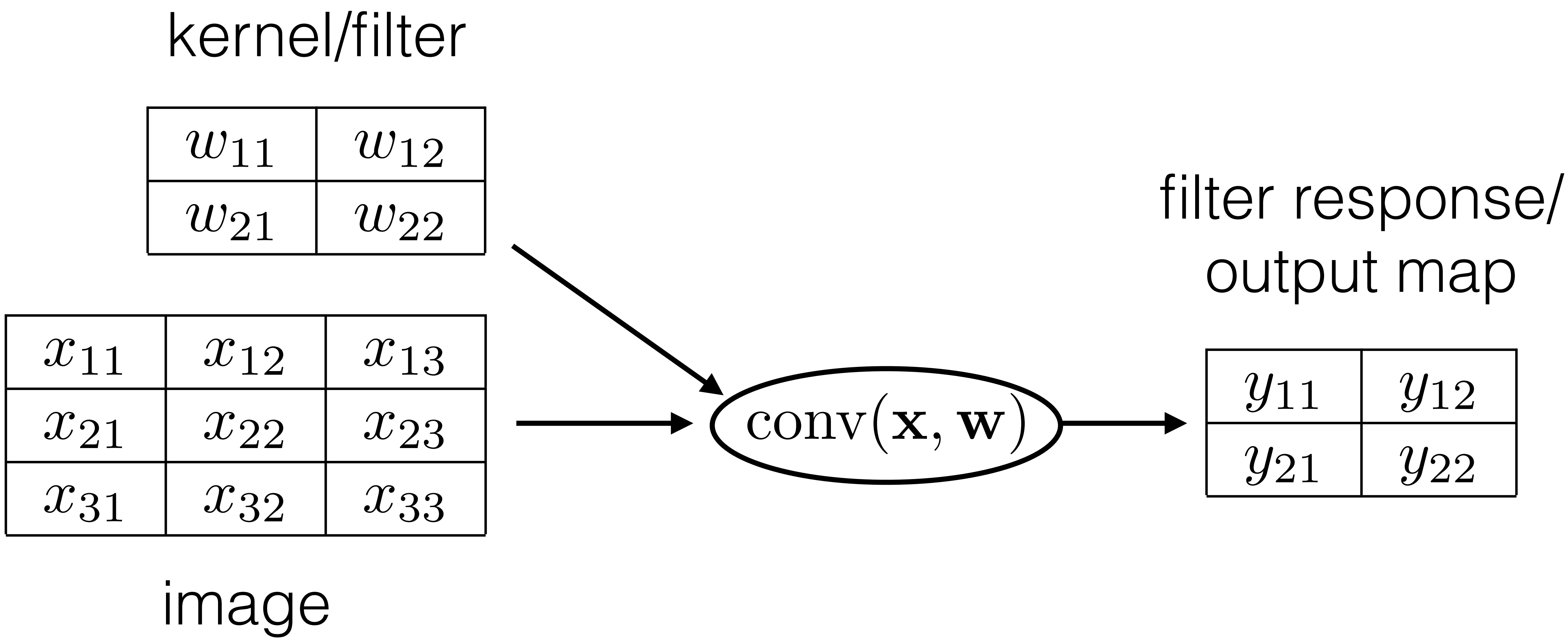


It corresponds to convolution of image $\mathbf{x}$ with kernel $\mathbf{w}$ followed by activation function



It corresponds to convolution of image $\mathbf{x}$ with kernel $\mathbf{w}$ followed by activation function

Convolution forward pass $\mathbf{y} = \mathrm{conv}(\mathbf{x}, \mathbf{w})$



Convolution forward pass $\mathbf{y} = \mathrm{conv}(\mathbf{x}, \mathbf{w})$

$$\begin{array}{|c|c|}\hline y_{11} & y_{12} \\ \hline y_{21} & y_{22} \\ \hline\end{array} = \mathrm{conv}\left(\begin{array}{|c|c|c|}\hline x_{11} & x_{12} & x_{13} \\ \hline x_{21} & x_{22} & x_{23} \\ \hline x_{31} & x_{32} & x_{33} \\ \hline\end{array} \, , \begin{array}{|c|c|}\hline w_{11} & w_{12} \\ \hline w_{21} & w_{22} \\ \hline\end{array} \right)$$

$$y_{11} = w_{11}x_{11} + w_{12}x_{12} + w_{21}x_{21} + w_{22}x_{22}$$

$$y_{12} = w_{11}x_{12} + w_{12}x_{13} + w_{21}x_{22} + w_{22}x_{23}$$

$$y_{21} = w_{11}x_{21} + w_{12}x_{22} + w_{21}x_{31} + w_{22}x_{32}$$

$$y_{22} = w_{11}x_{22} + w_{12}x_{23} + w_{21}x_{32} + w_{22}x_{33}$$

Convolution forward pass $\mathbf{y} = \mathrm{conv}(\mathbf{x}, \mathbf{w})$

$$\begin{array}{|c|c|}\hline y_{11} & y_{12} \\ \hline y_{21} & y_{22} \\ \hline \end{array} = \mathrm{conv}\left(\begin{array}{|c|c|c|}\hline x_{11} & x_{12} & x_{13} \\ \hline x_{21} & x_{22} & x_{23} \\ \hline x_{31} & x_{32} & x_{33} \\ \hline \end{array}, \begin{array}{|c|c|}\hline w_{11} & w_{12} \\ \hline w_{21} & w_{22} \\ \hline \end{array} \right)$$

$$y_{11} = w_{11}x_{11} + w_{12}x_{12} + w_{21}x_{21} + w_{22}x_{22}$$

$$y_{12} = w_{11}x_{12} + w_{12}x_{13} + w_{21}x_{22} + w_{22}x_{23}$$

$$y_{21} = w_{11}x_{21} + w_{12}x_{22} + w_{21}x_{31} + w_{22}x_{32}$$

$$y_{22} = w_{11}x_{22} + w_{12}x_{23} + w_{21}x_{32} + w_{22}x_{33}$$

Convolution forward pass $\mathbf{y} = \mathrm{conv}(\mathbf{x}, \mathbf{w})$

$$\begin{array}{|c|c|}\hline y_{11} & \boxed{y_{12}} \\ \hline y_{21} & y_{22} \\ \hline \end{array} = \mathrm{conv}\left(\begin{array}{|c|c|c|}\hline x_{11} & x_{12} & x_{13} \\ \hline x_{21} & x_{22} & x_{23} \\ \hline x_{31} & x_{32} & x_{33} \\ \hline \end{array}, \begin{array}{|c|c|}\hline w_{11} & w_{12} \\ \hline w_{21} & w_{22} \\ \hline \end{array} \right)$$

$$y_{11} = w_{11}x_{11} + w_{12}x_{12} + w_{21}x_{21} + w_{22}x_{22}$$

$$\boxed{y_{12} = w_{11}x_{12} + w_{12}x_{13} + w_{21}x_{22} + w_{22}x_{23}}$$

$$y_{21} = w_{11}x_{21} + w_{12}x_{22} + w_{21}x_{31} + w_{22}x_{32}$$

$$y_{22} = w_{11}x_{22} + w_{12}x_{23} + w_{21}x_{32} + w_{22}x_{33}$$

Convolution forward pass $\mathbf{y} = \mathrm{conv}(\mathbf{x}, \mathbf{w})$

$$\begin{array}{|c|c|}\hline y_{11} & y_{12} \\ \hline y_{21} & y_{22} \\ \hline\end{array} = \mathrm{conv}\left(\begin{array}{|c|c|c|}\hline x_{11} & x_{12} & x_{13} \\ \hline x_{21} & x_{22} & x_{23} \\ \hline x_{31} & x_{32} & x_{33} \\ \hline\end{array}, \begin{array}{|c|c|}\hline w_{11} & w_{12} \\ \hline w_{21} & w_{22} \\ \hline\end{array} \right)$$

$$y_{11} = w_{11}x_{11} + w_{12}x_{12} + w_{21}x_{21} + w_{22}x_{22}$$

$$y_{12} = w_{11}x_{12} + w_{12}x_{13} + w_{21}x_{22} + w_{22}x_{23}$$

$$y_{21} = w_{11}x_{21} + w_{12}x_{22} + w_{21}x_{31} + w_{22}x_{32}$$

$$y_{22} = w_{11}x_{22} + w_{12}x_{23} + w_{21}x_{32} + w_{22}x_{33}$$

Convolution forward pass $\mathbf{y} = \mathrm{conv}(\mathbf{x}, \mathbf{w})$

$$\begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix} = \mathrm{conv}\left(\begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{bmatrix}, \begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{bmatrix} \right)$$

$$y_{11} = w_{11}x_{11} + w_{12}x_{12} + w_{21}x_{21} + w_{22}x_{22}$$

$$y_{12} = w_{11}x_{12} + w_{12}x_{13} + w_{21}x_{22} + w_{22}x_{23}$$

$$y_{21} = w_{11}x_{21} + w_{12}x_{22} + w_{21}x_{31} + w_{22}x_{32}$$

$$y_{22} = w_{11}x_{22} + w_{12}x_{23} + w_{21}x_{32} + w_{22}x_{33}$$

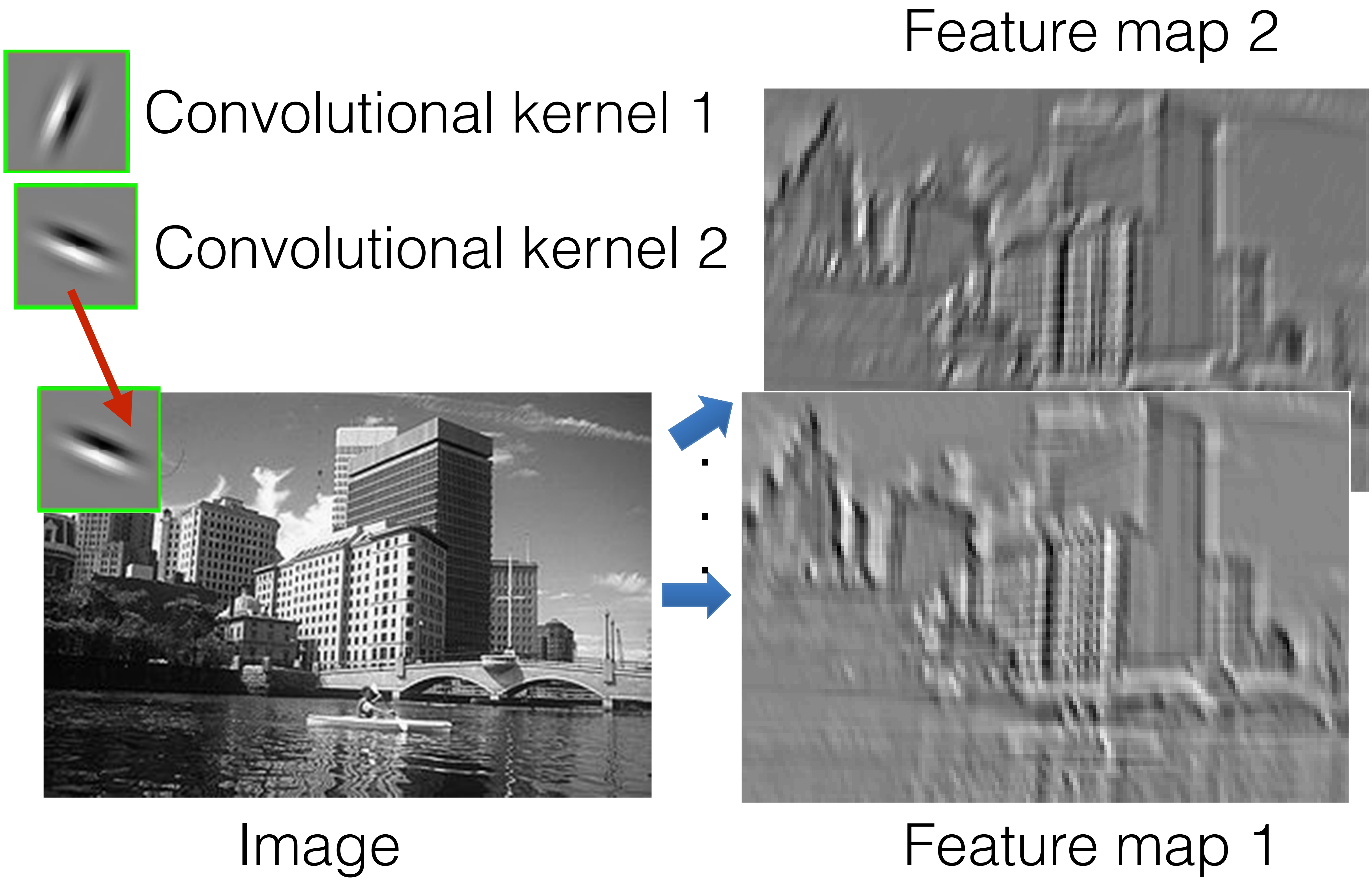
Feature maps

Convolutional kernel 1

Image

Feature map 1

Feature maps



Convolution layer properties - output size

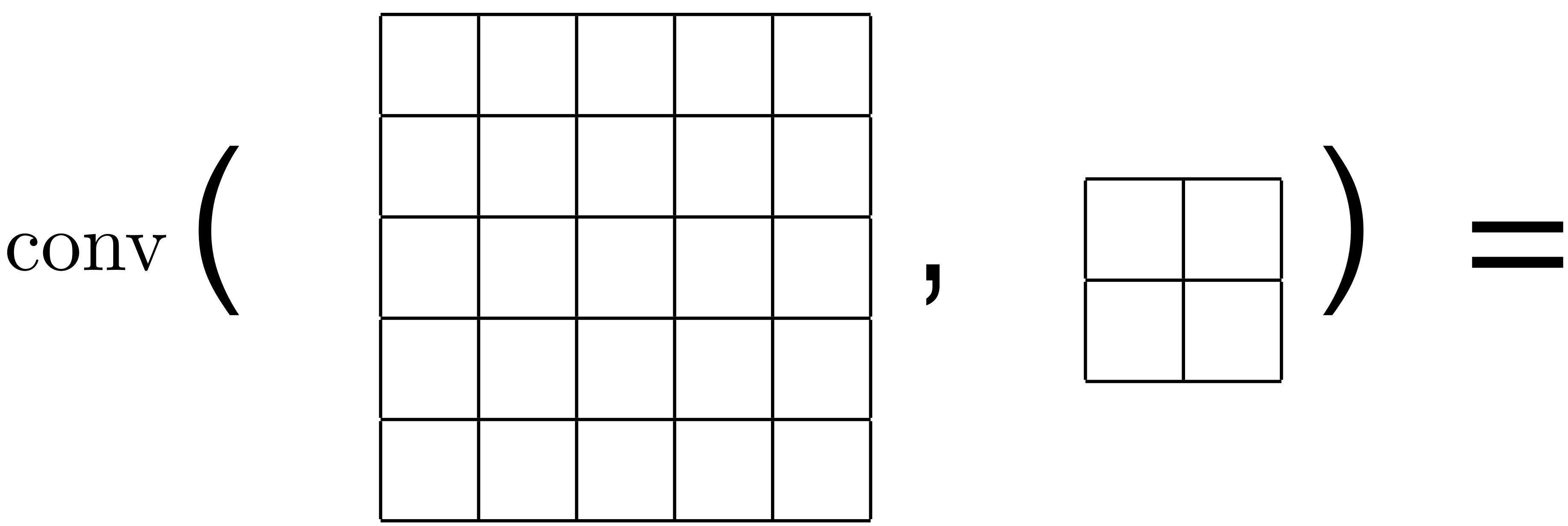


image (5x5) | kernel (2x2) | output (**? x ?**)

Convolution layer properties - output size

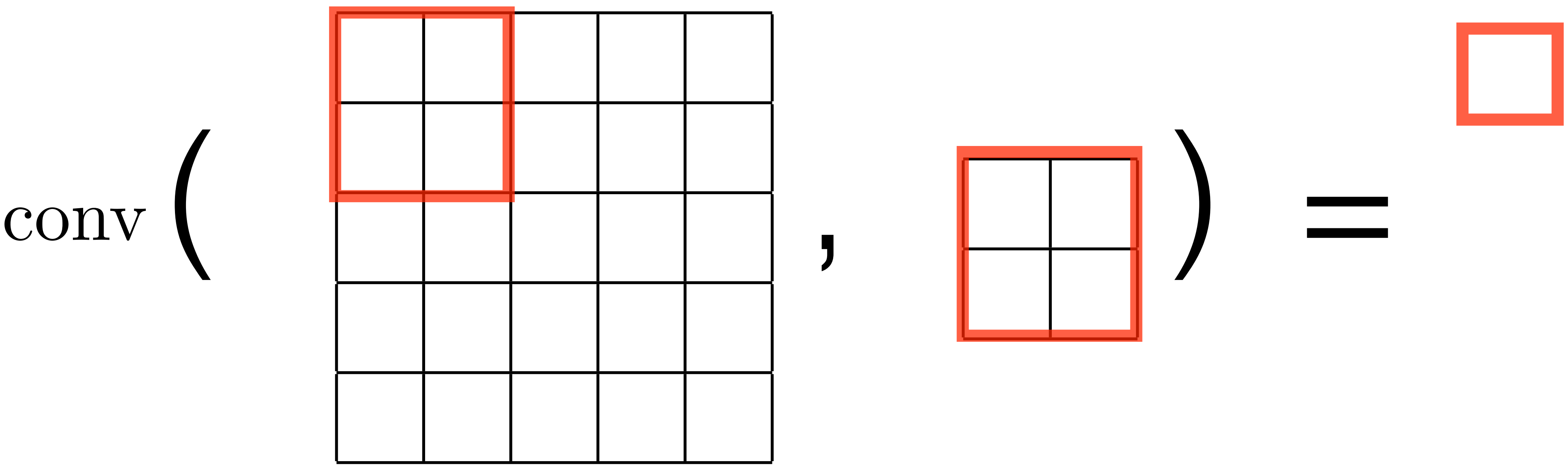


image (5x5) | kernel (2x2) | output (**? x ?**)

Convolution layer properties - output size

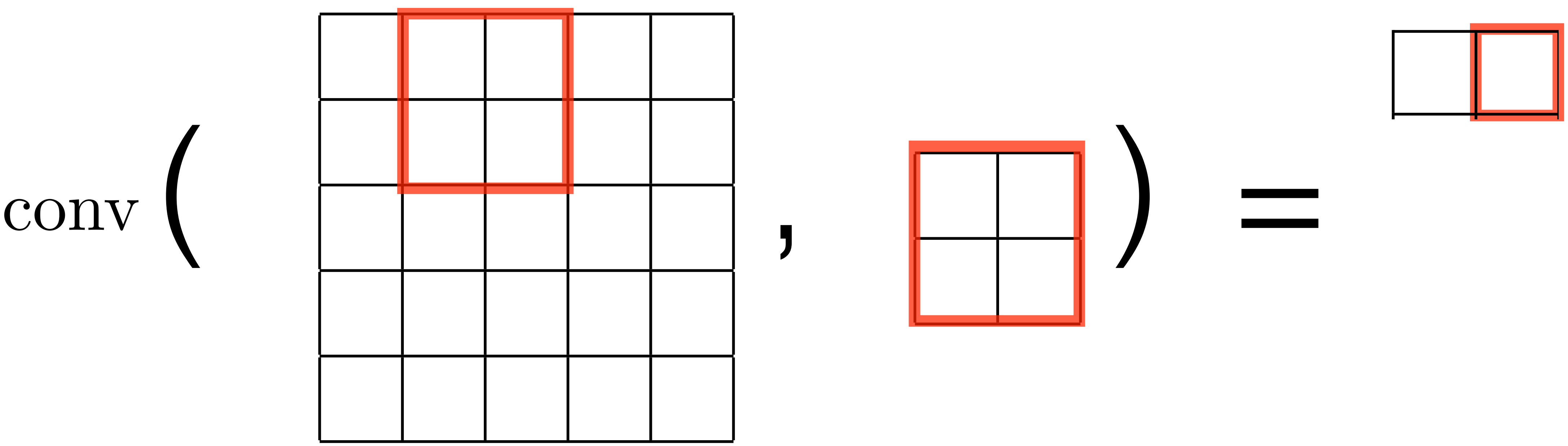


image (5x5)

kernel (2x2)

output (**? x ?**)

Convolution layer properties - output size

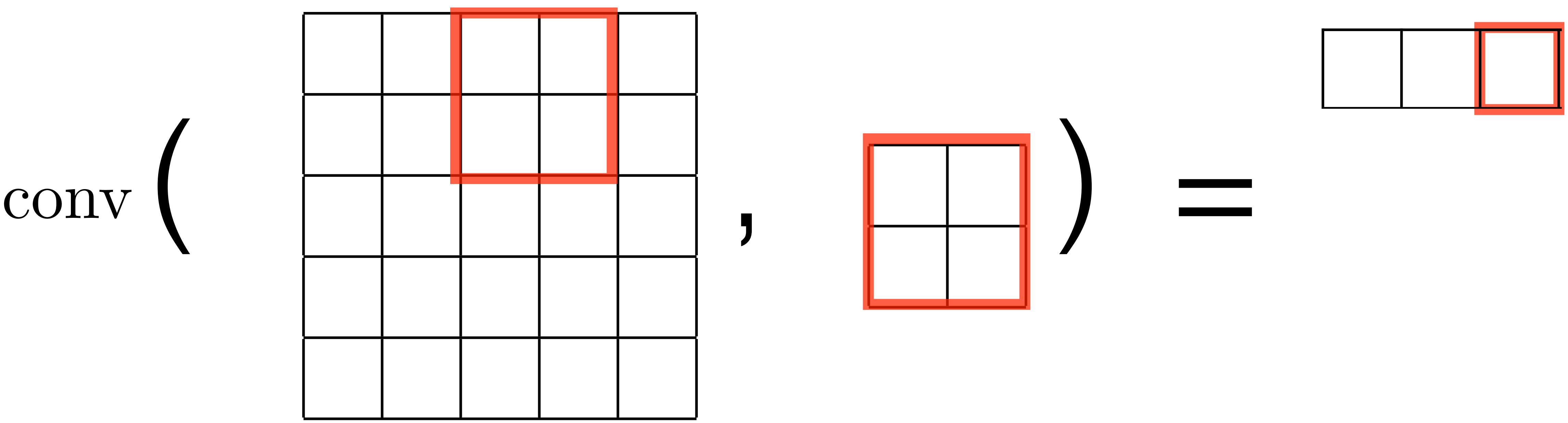


image
(5x5)

kernel
(2x2)

output
(**? x ?**)

Convolution layer properties - output size

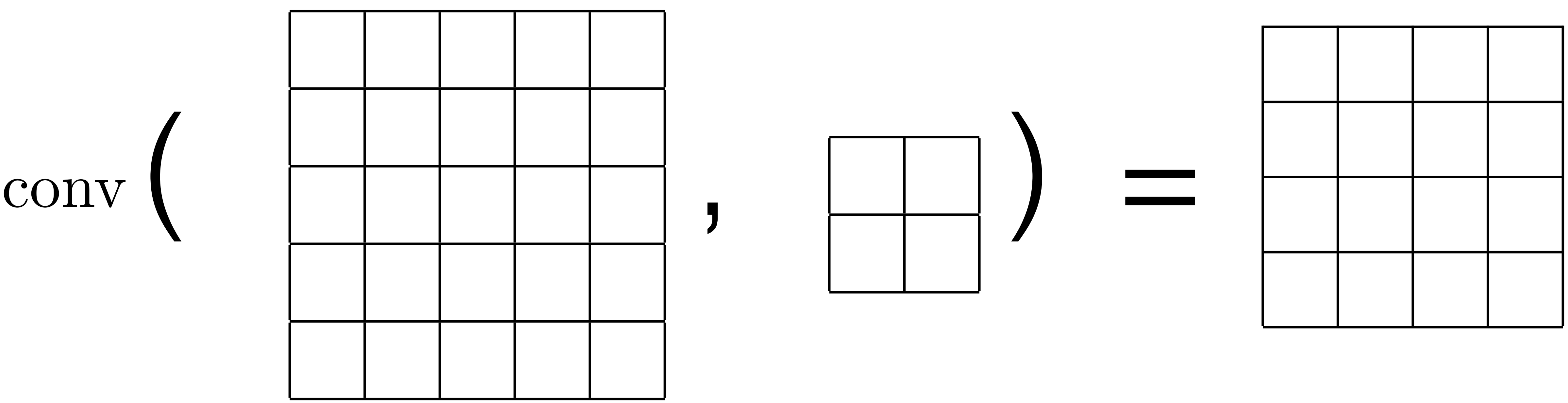


image (5x5)

kernel (2x2)

output (4x4)

Convolution layer properties - output size

M = N - K + 1

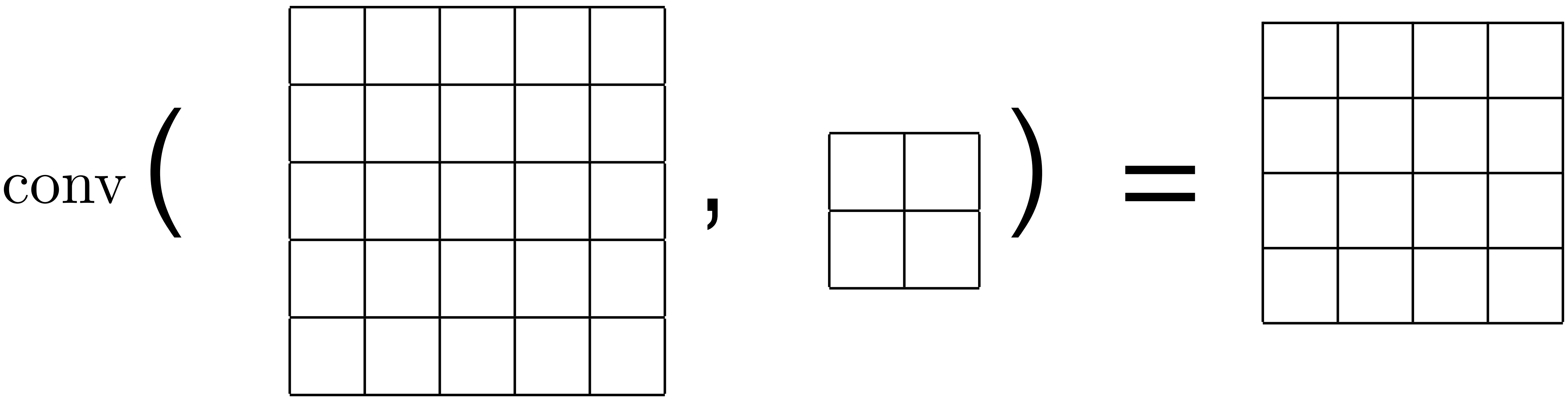


image (NxN)

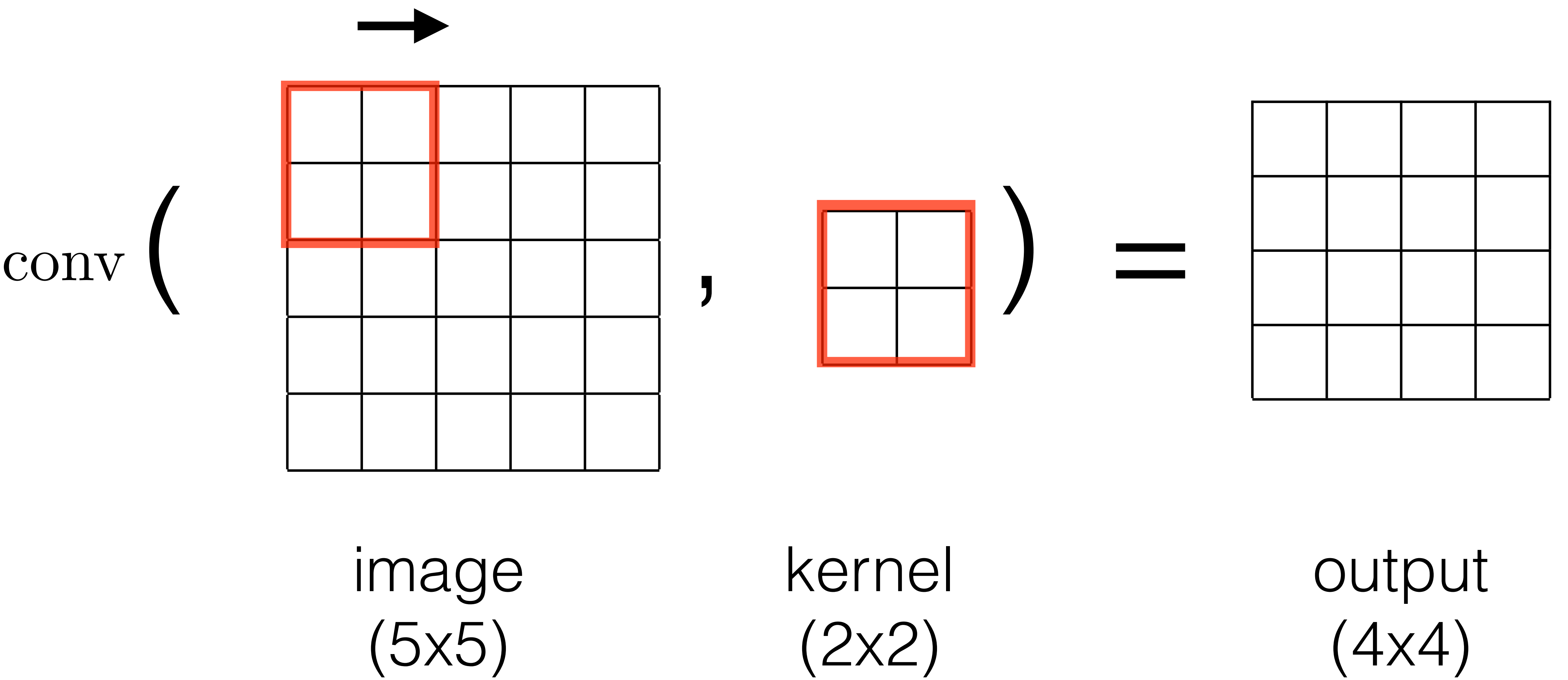
kernel (KxK)

output (MxM)

Convolution layer properties - stride

stride = 1

kernel moves by 1 pixel

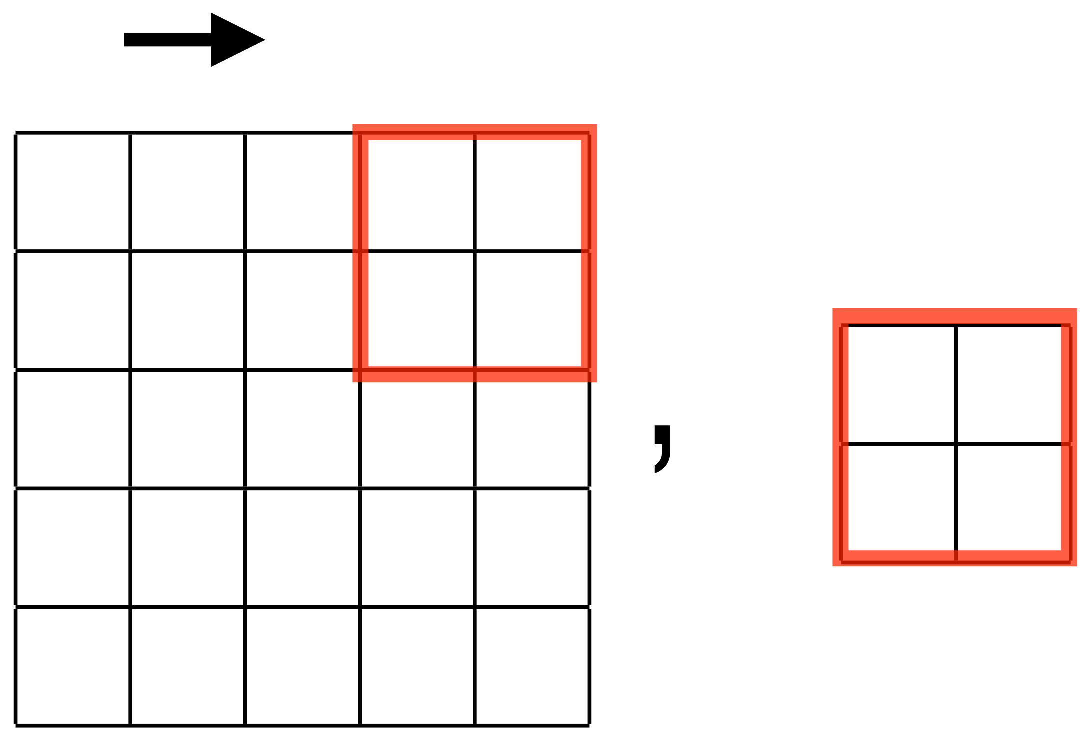


Convolution layer properties - stride

stride = 3

kernel moves by 3 pixels

conv (



) =

image
(5x5)

kernel
(2x2)

output
(**? x ?**)

Convolution layer properties - stride

stride = 3

kernel moves by 3 pixels

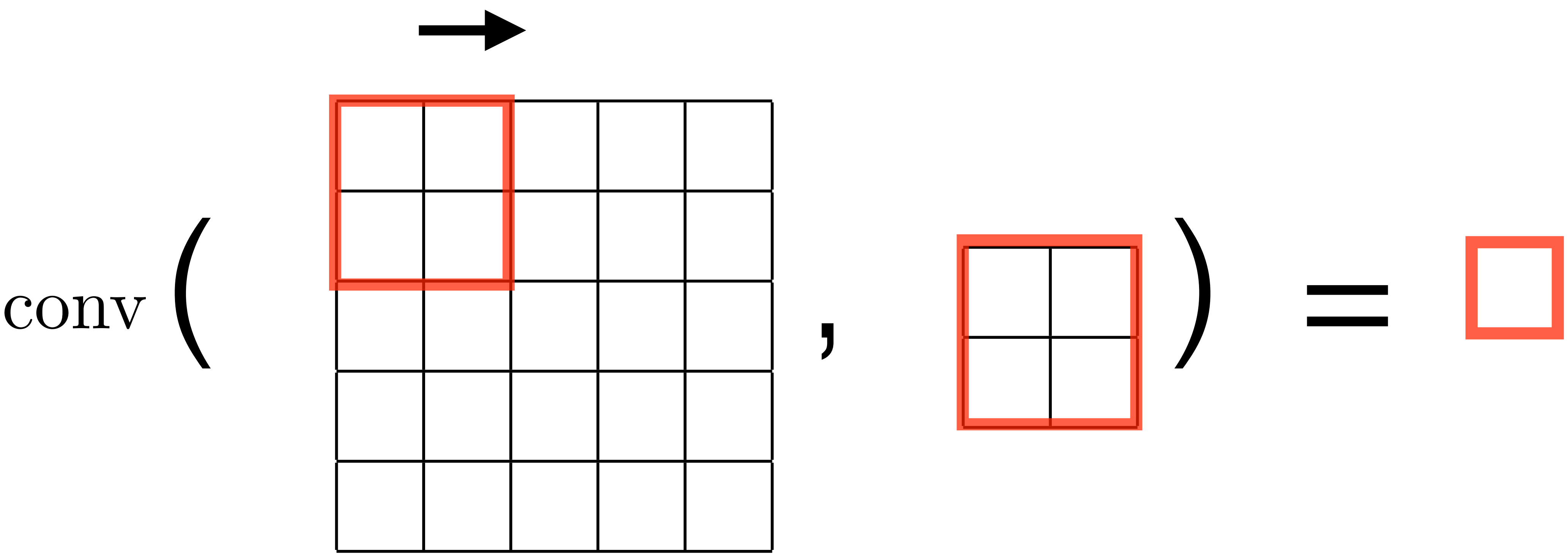


image (5x5)

kernel (2x2)

output (? x ?)

Convolution layer properties - stride

stride = 3

kernel moves by 3 pixels

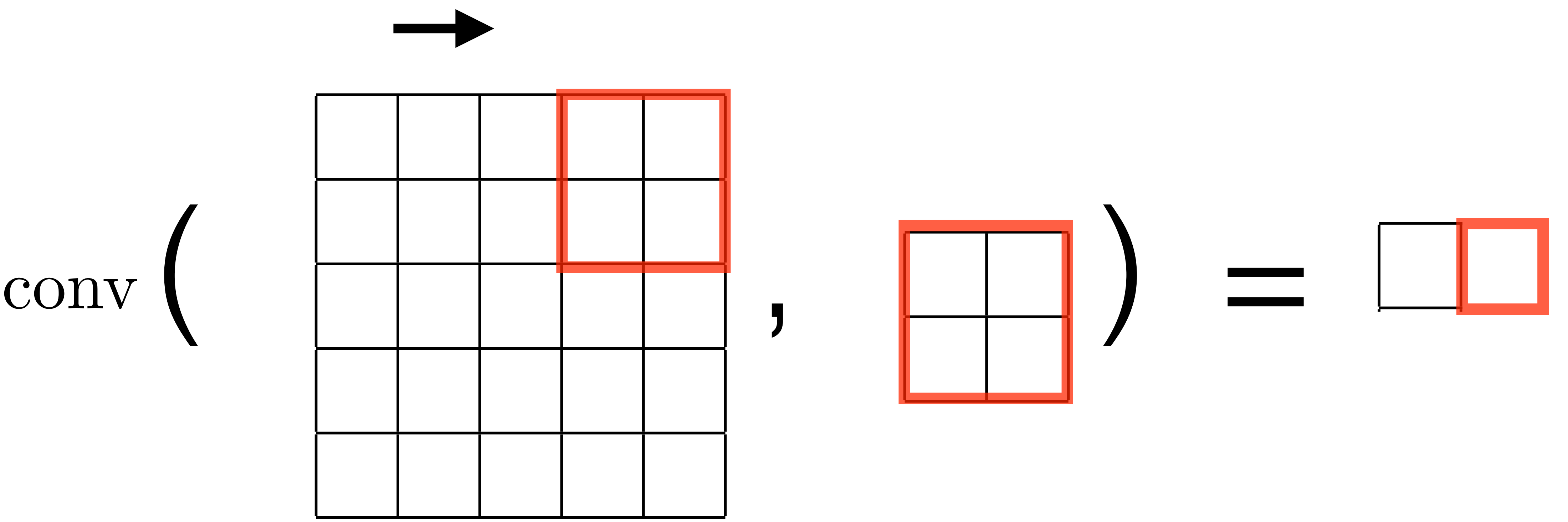


image
(5x5)

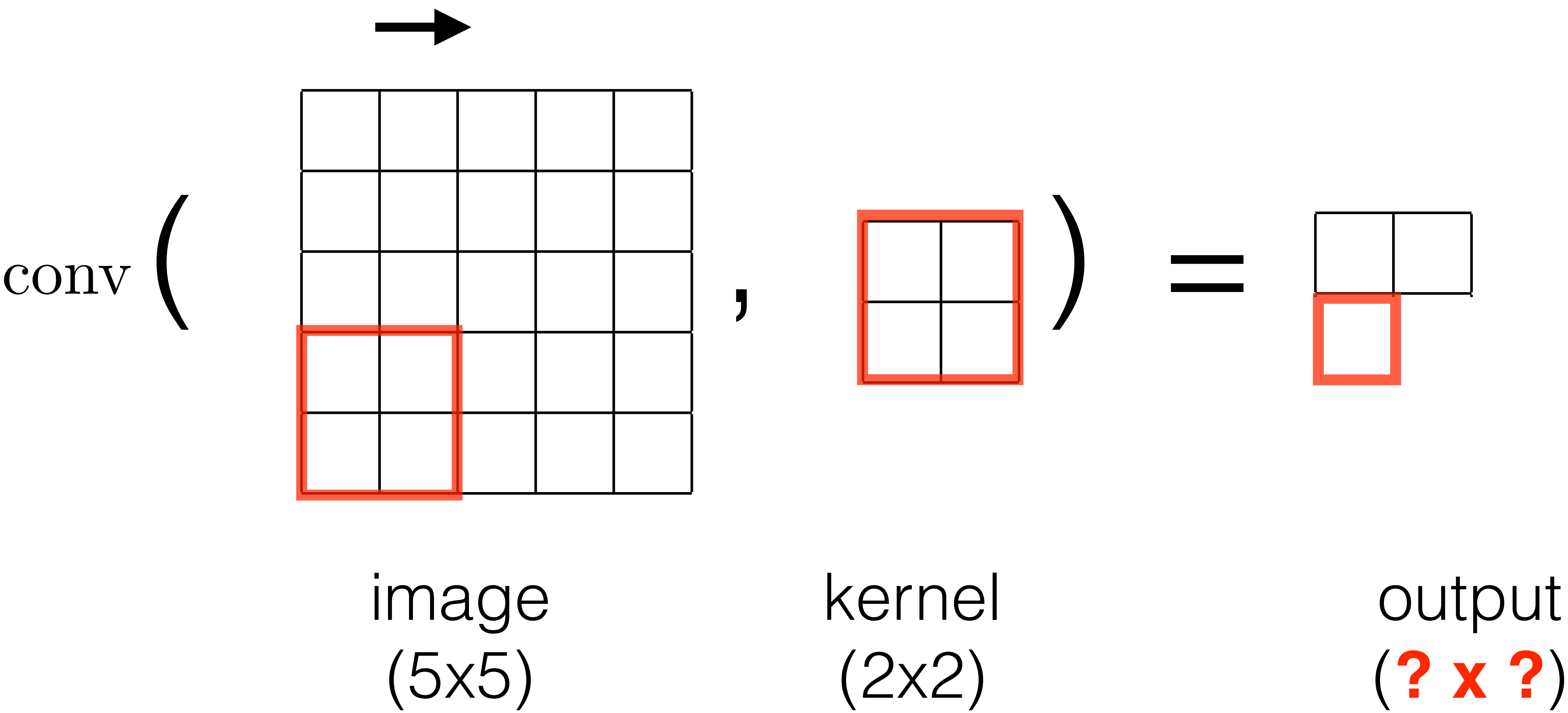
kernel
(2x2)

output
(**? x ?**)

Convolution layer properties - stride

stride = 3

kernel moves by 3 pixels



Convolution layer properties - stride

stride = 3

kernel moves by 3 pixels

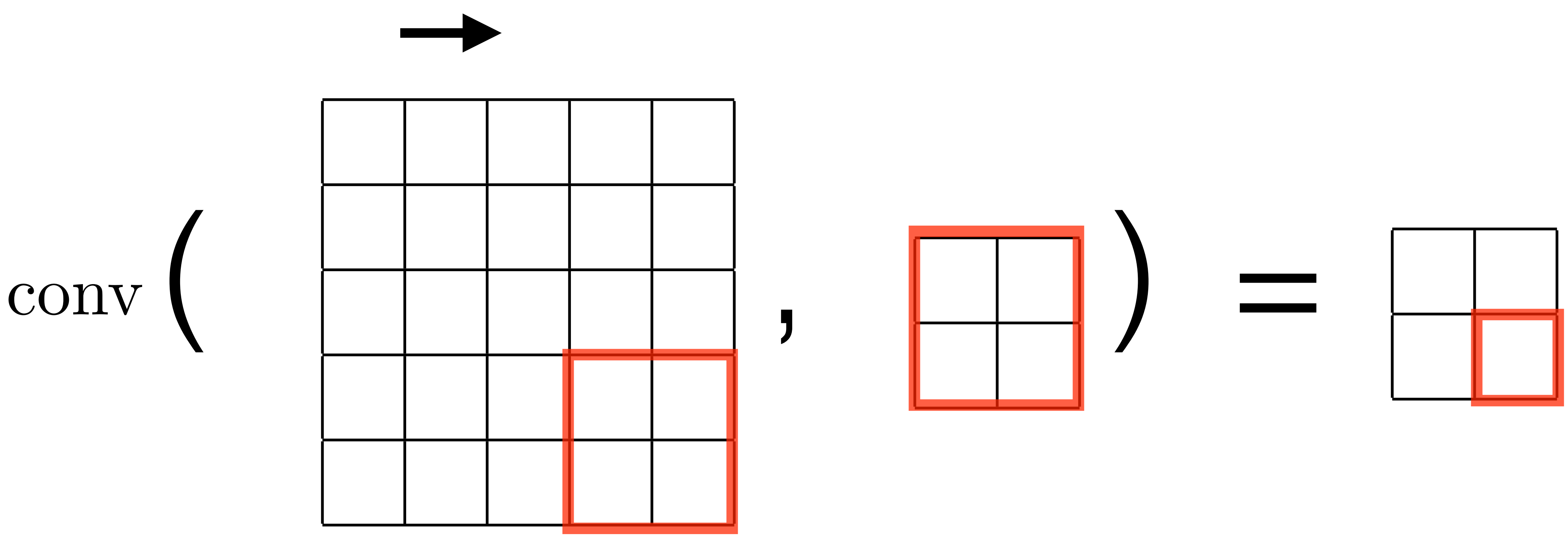


image (5x5)

kernel (2x2)

output (2x2)

Convolution layer properties - stride

M = **floor**((N-K) / stride + 1)

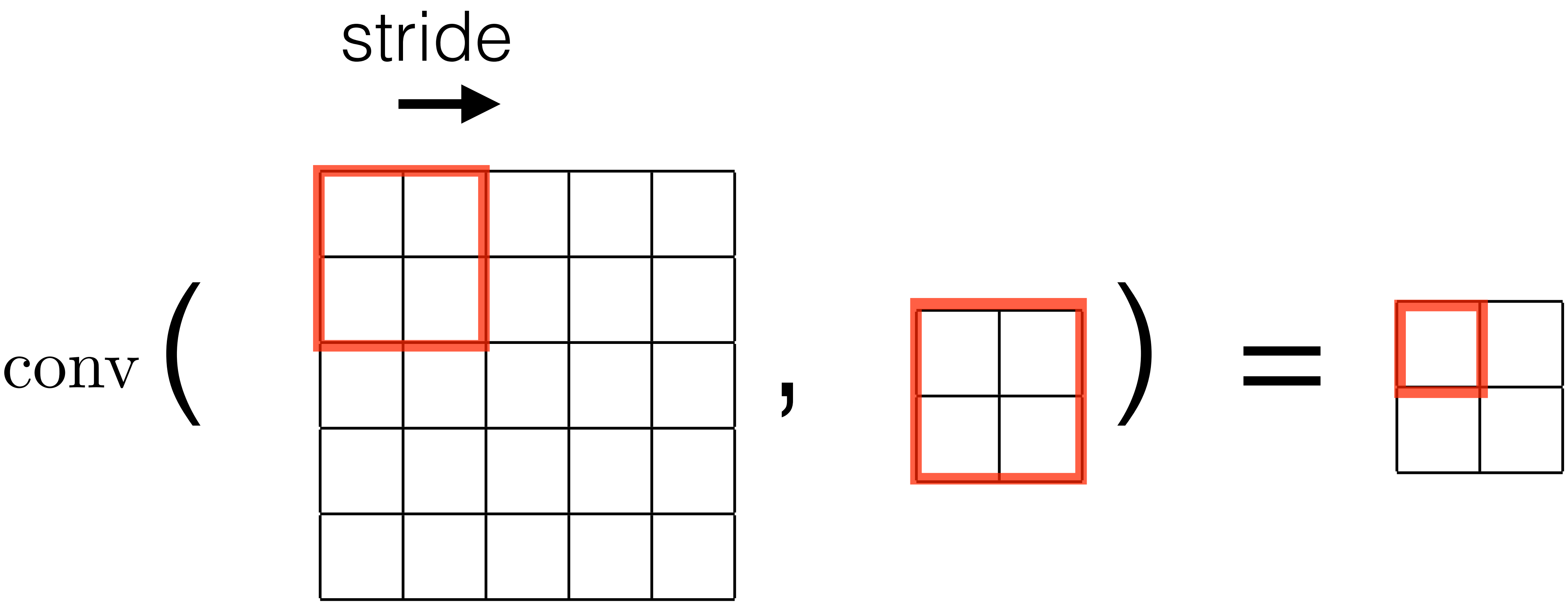
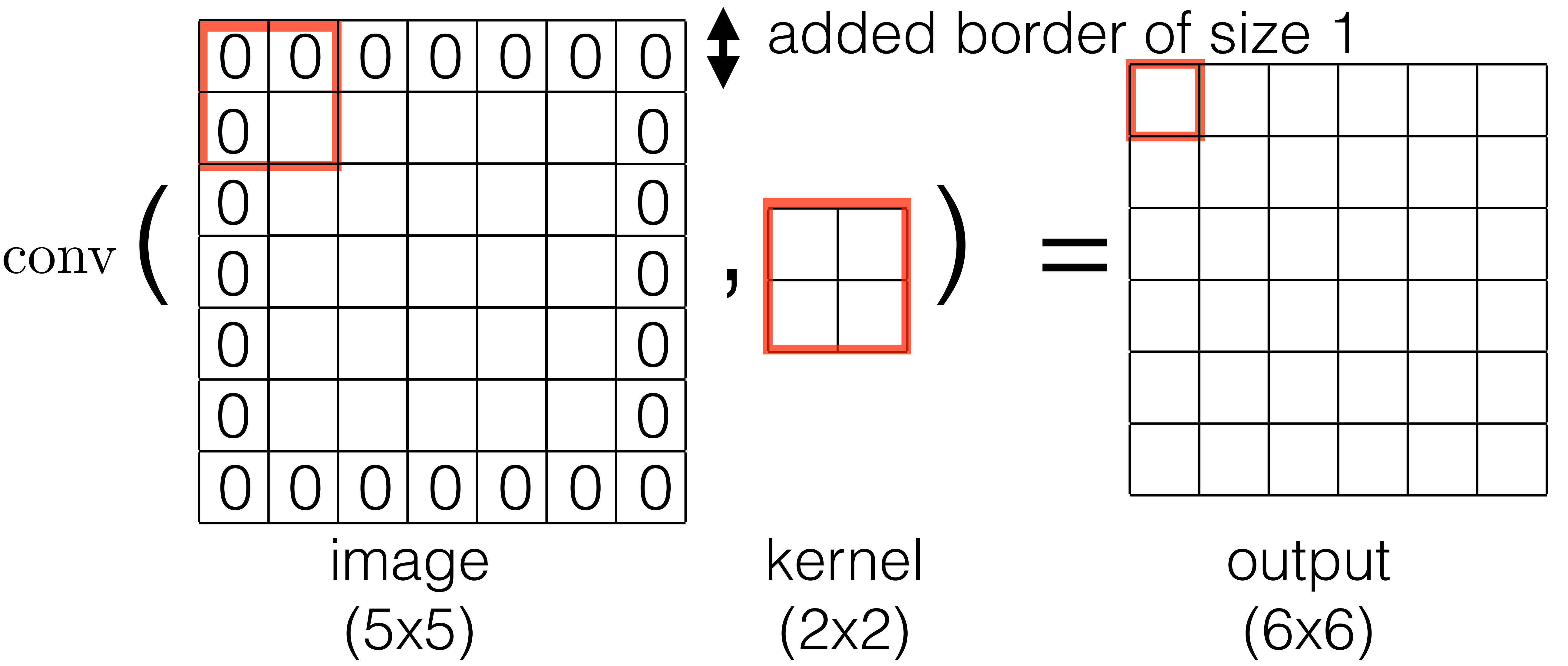


image (NxN) | kernel (KxK) | output (MxM)

e.g. M = (5-2) / 3 +1 = 2

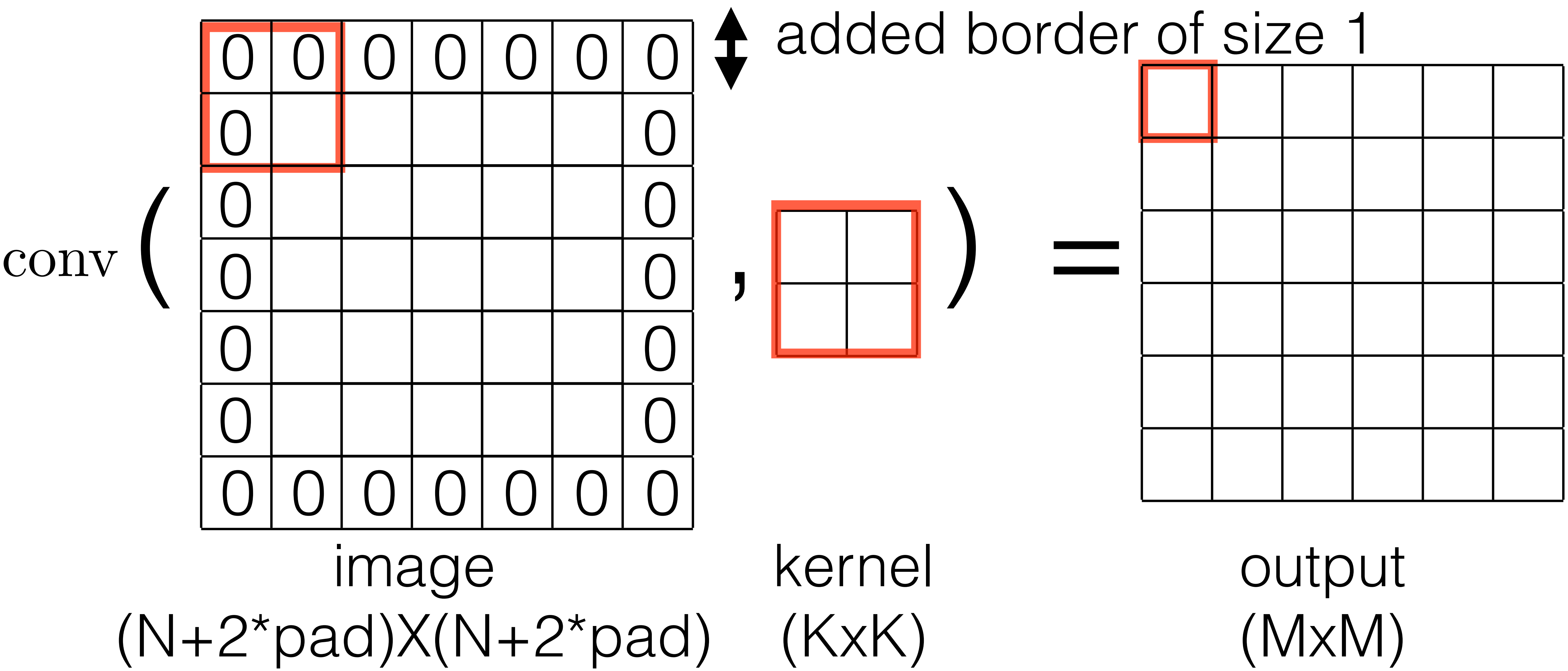
Convolution layer properties - pad

pad = 1



Convolution layer properties - pad

$$M = \texttt{floor}\left((N+2^*pad-K) / stride + 1\right)$$



Convolution layer

Dilatation rate = 1

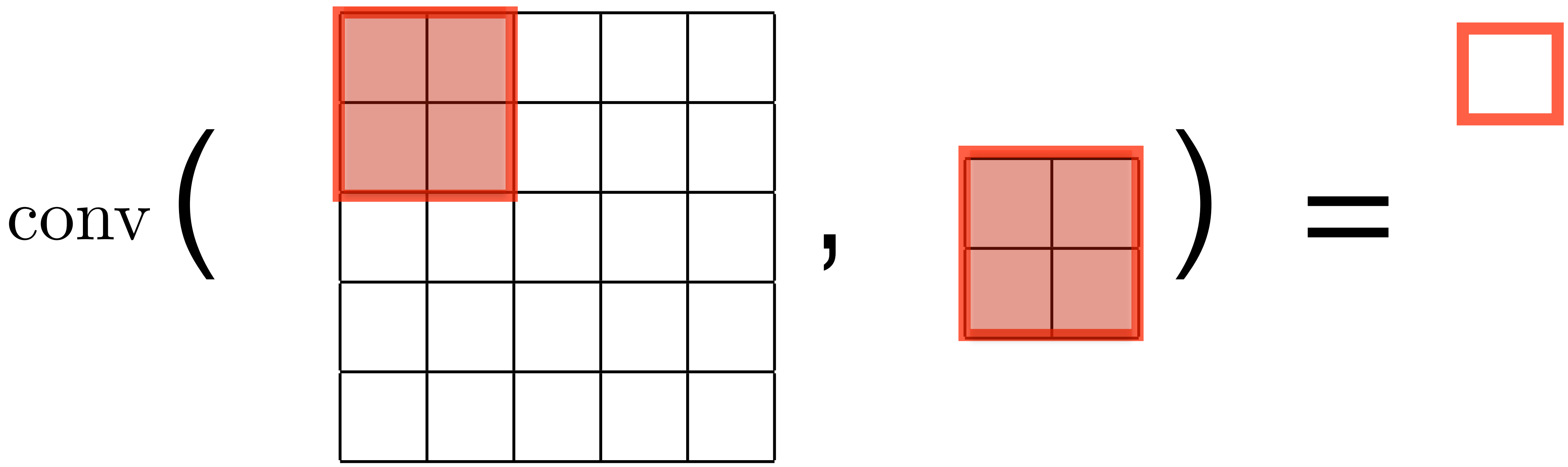


image (5x5) | kernel (2x2) | output (**? x ?**)

Atrous convolution layer

Dilatation rate = 2

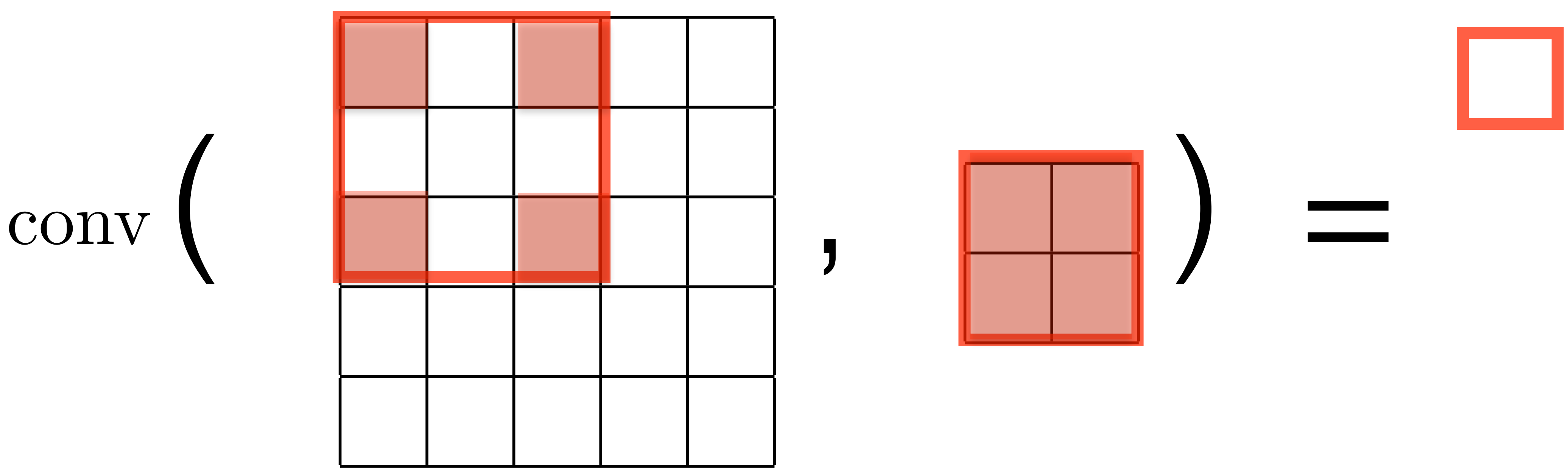
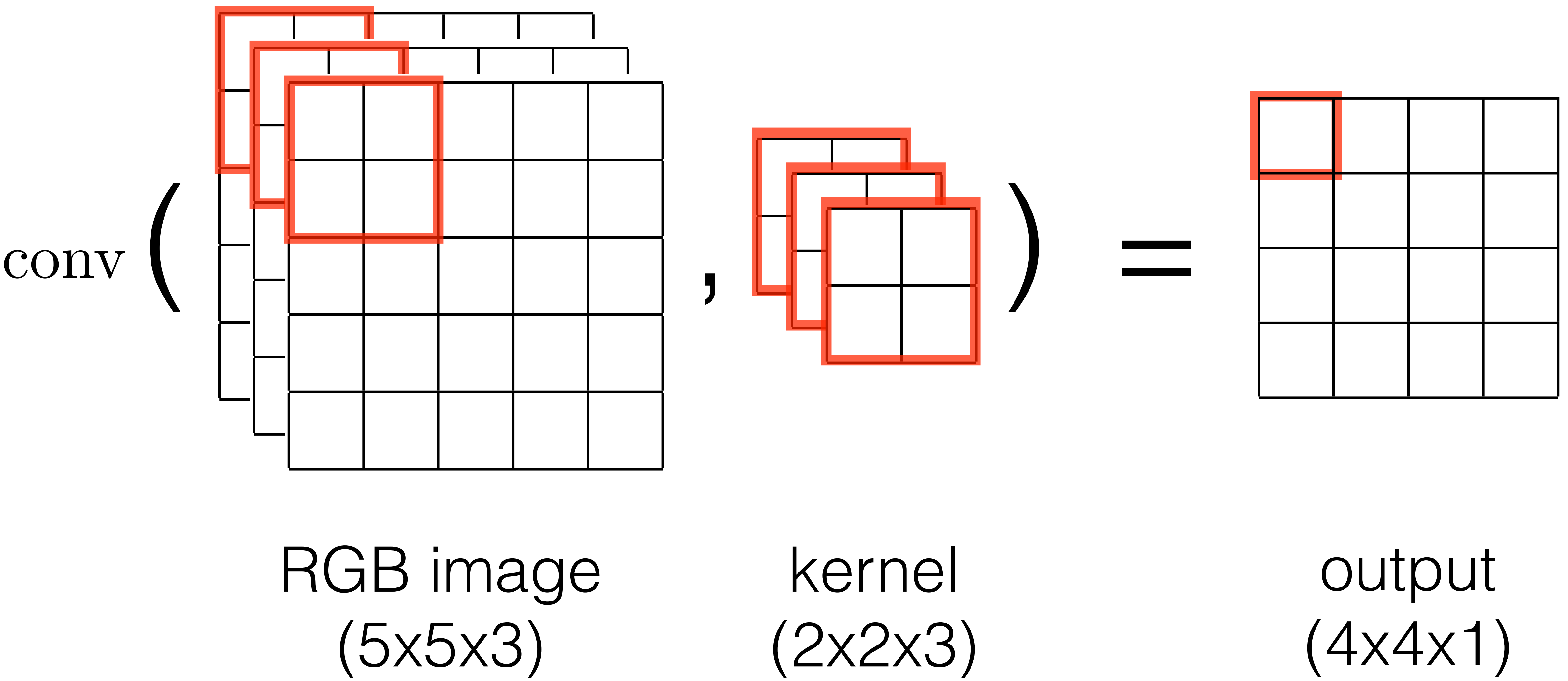


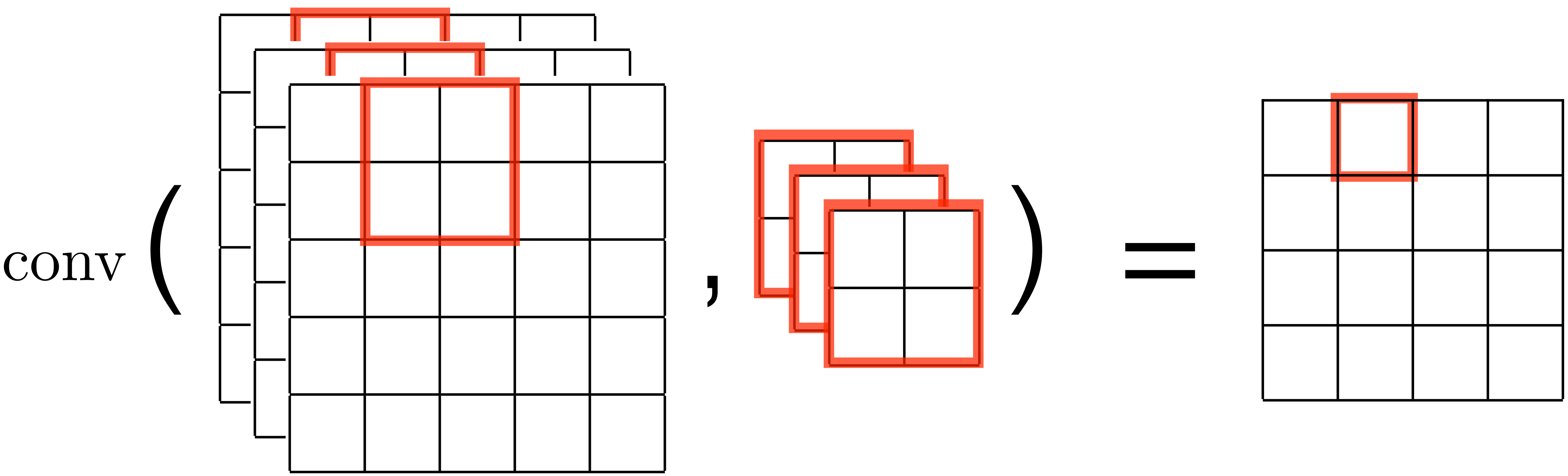
image (5x5) kernel (2x2) output (**? x ?**)

Show python code

Multi-channel convolution



Multi-channel convolution

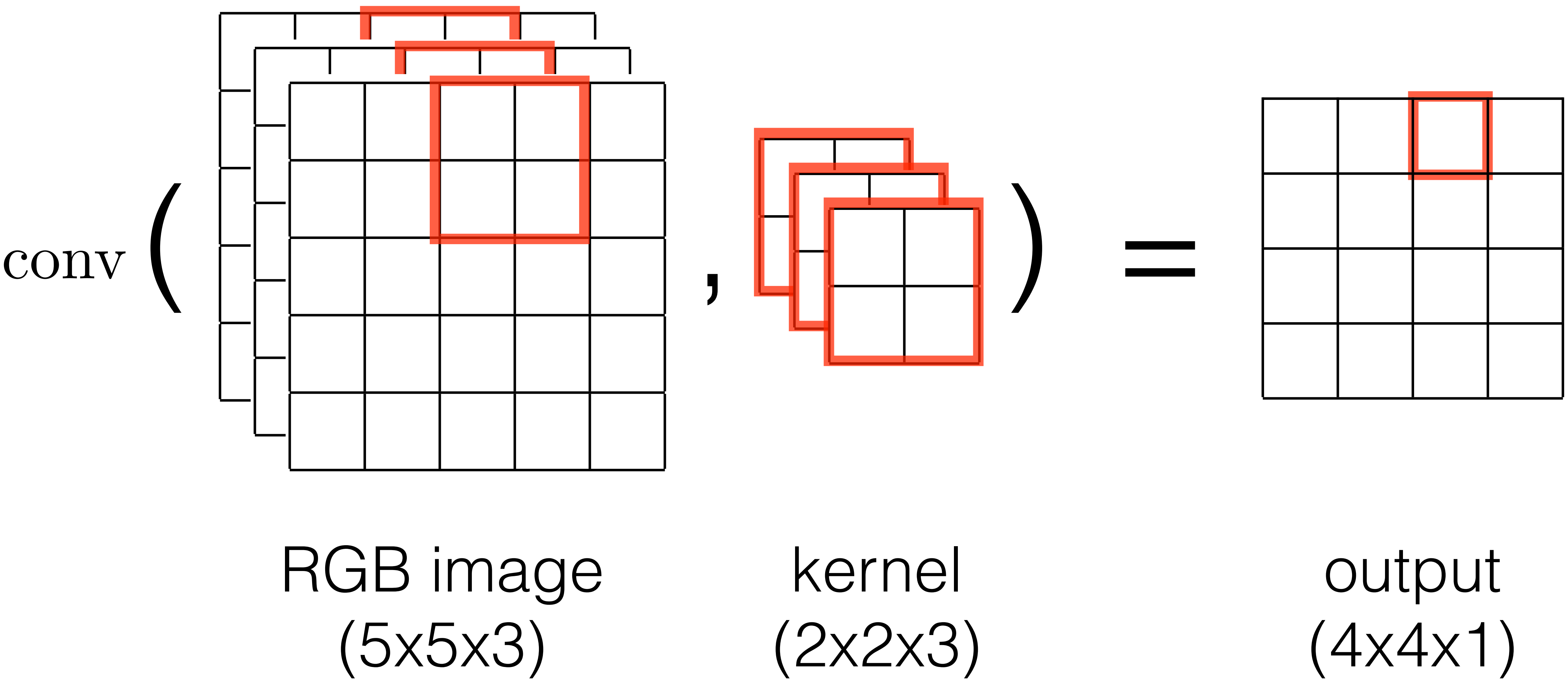


RGB image
(5x5x3)

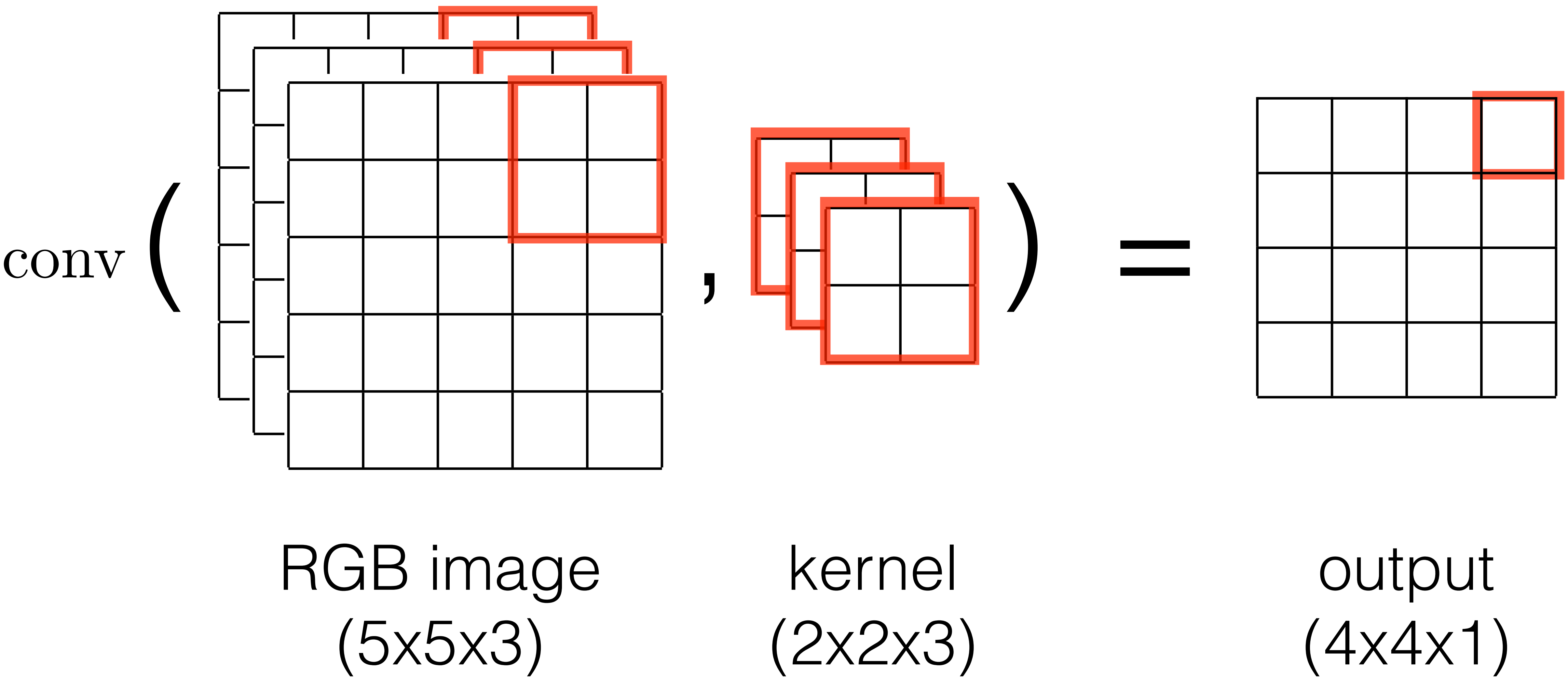
kernel
(2x2x3)

output
(4x4x1)

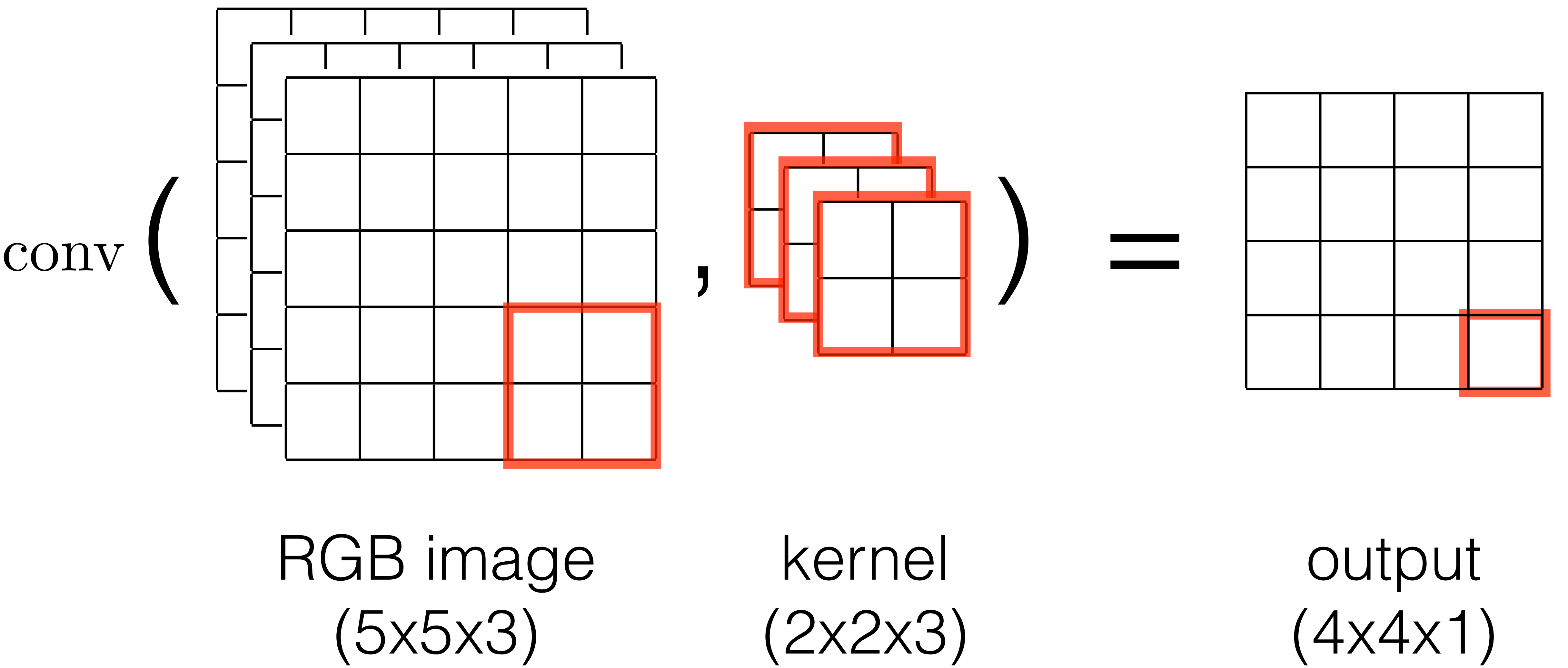
Multi-channel convolution



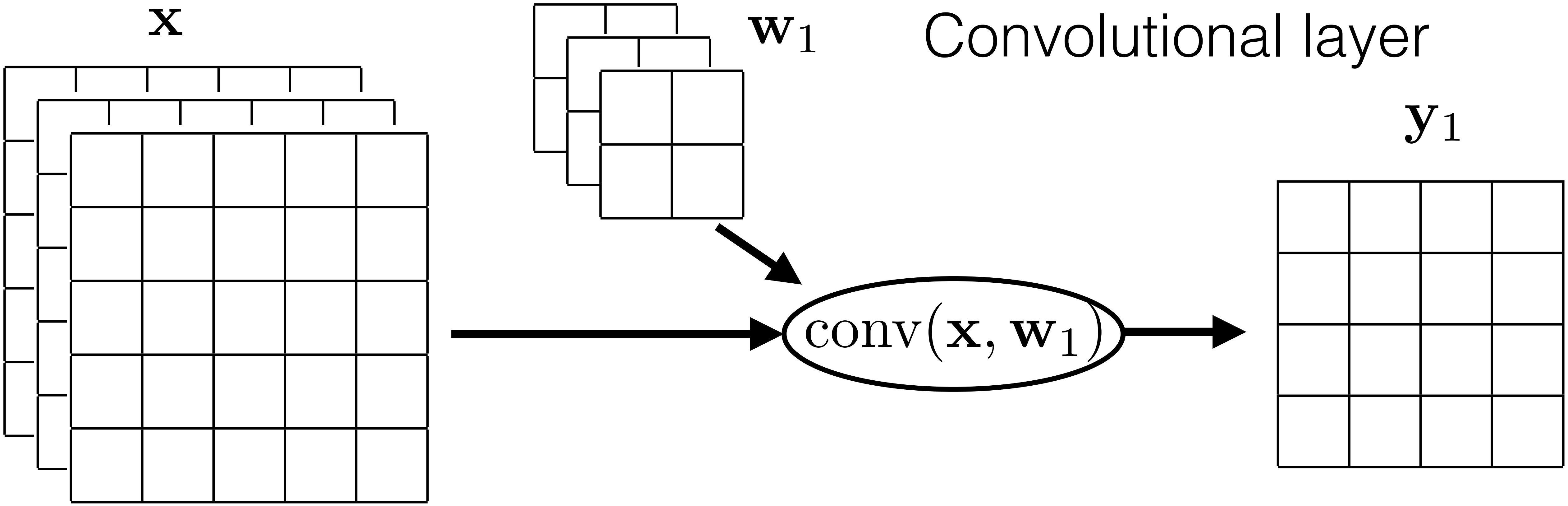
Multi-channel convolution



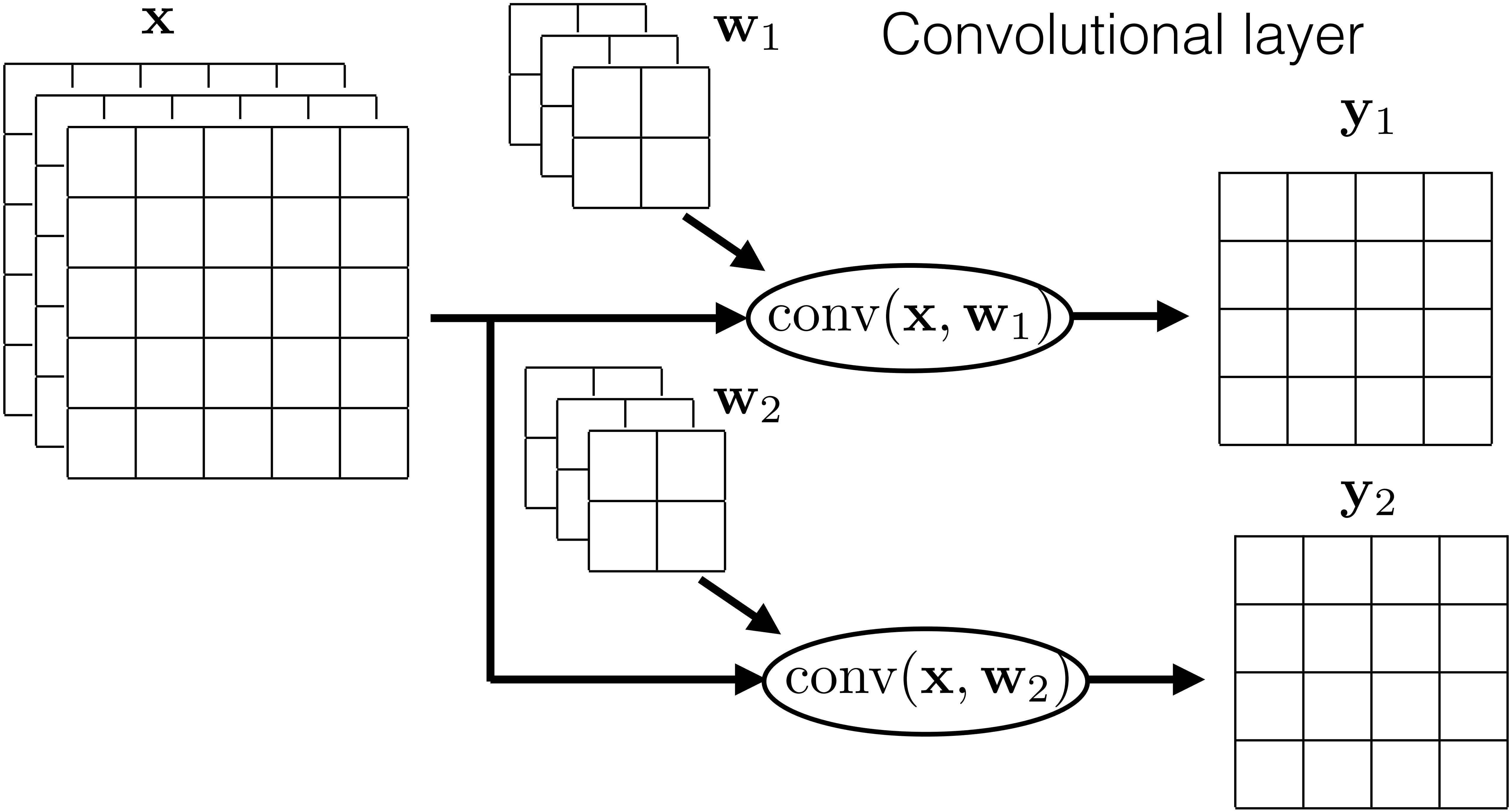
Multi-channel convolution



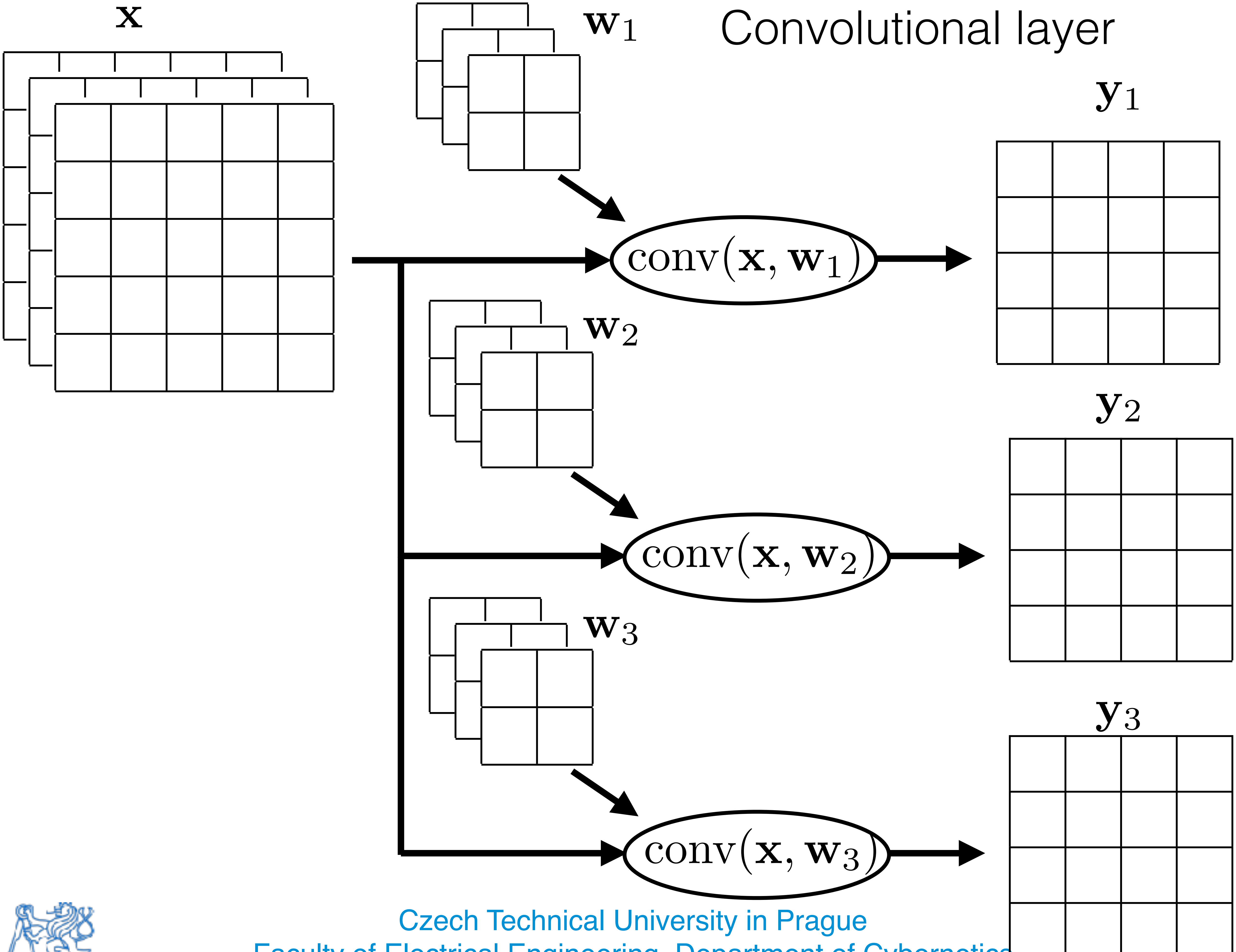
Convolutional layer



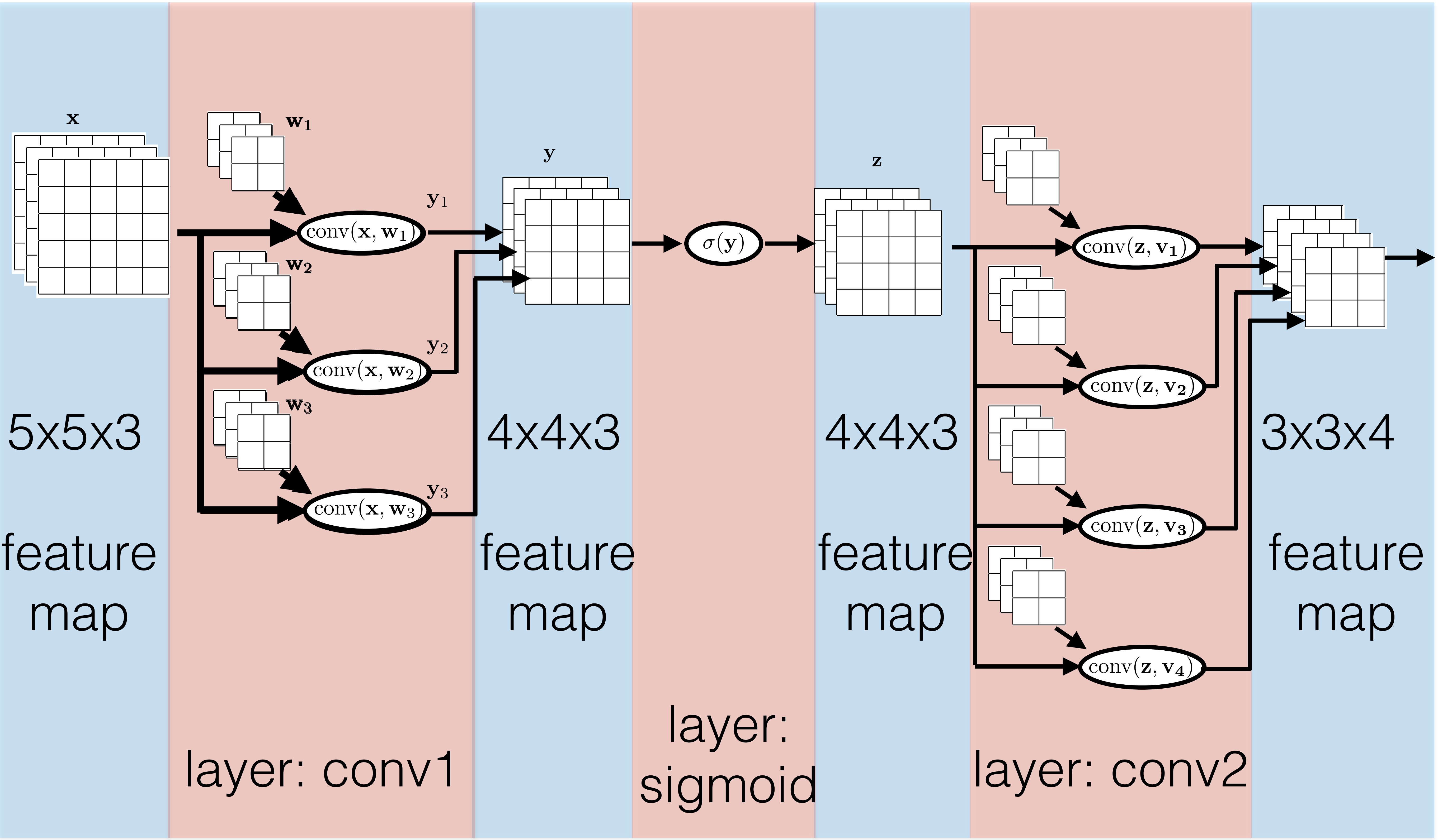
Convolutional layer



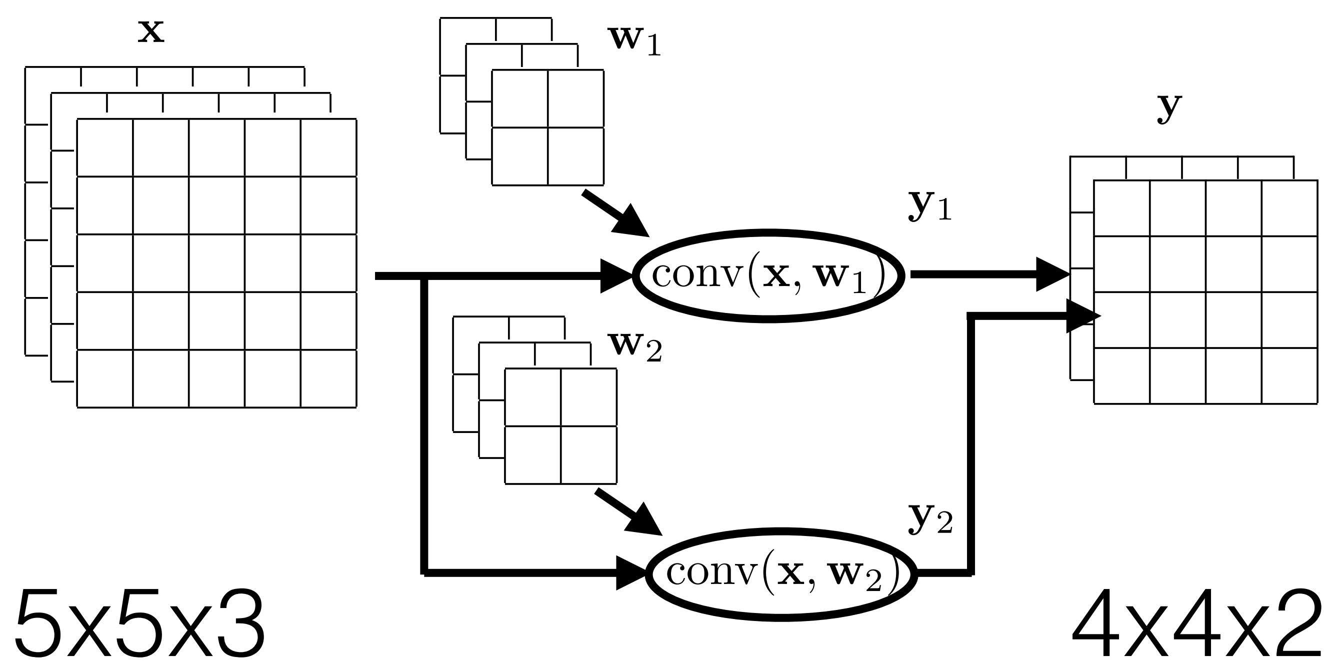
Convolutional layer



Convolutional network (ConvNet)

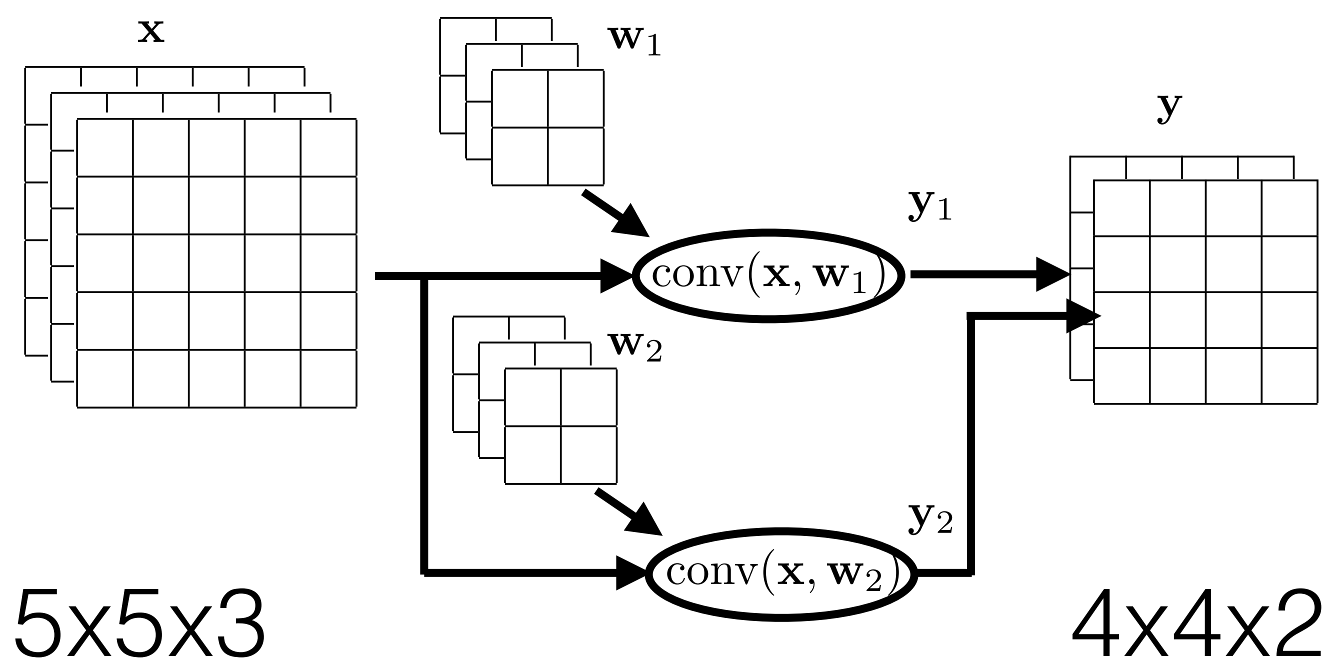


2D convolution forward pass




```
# initialise
import torch.nn as nn
# define 2D convolutional layer
first_layer = nn.Conv2d(in_channels=3, out_channels=2,
                        kernel_size=2, stride=1,
                        padding=1)
```

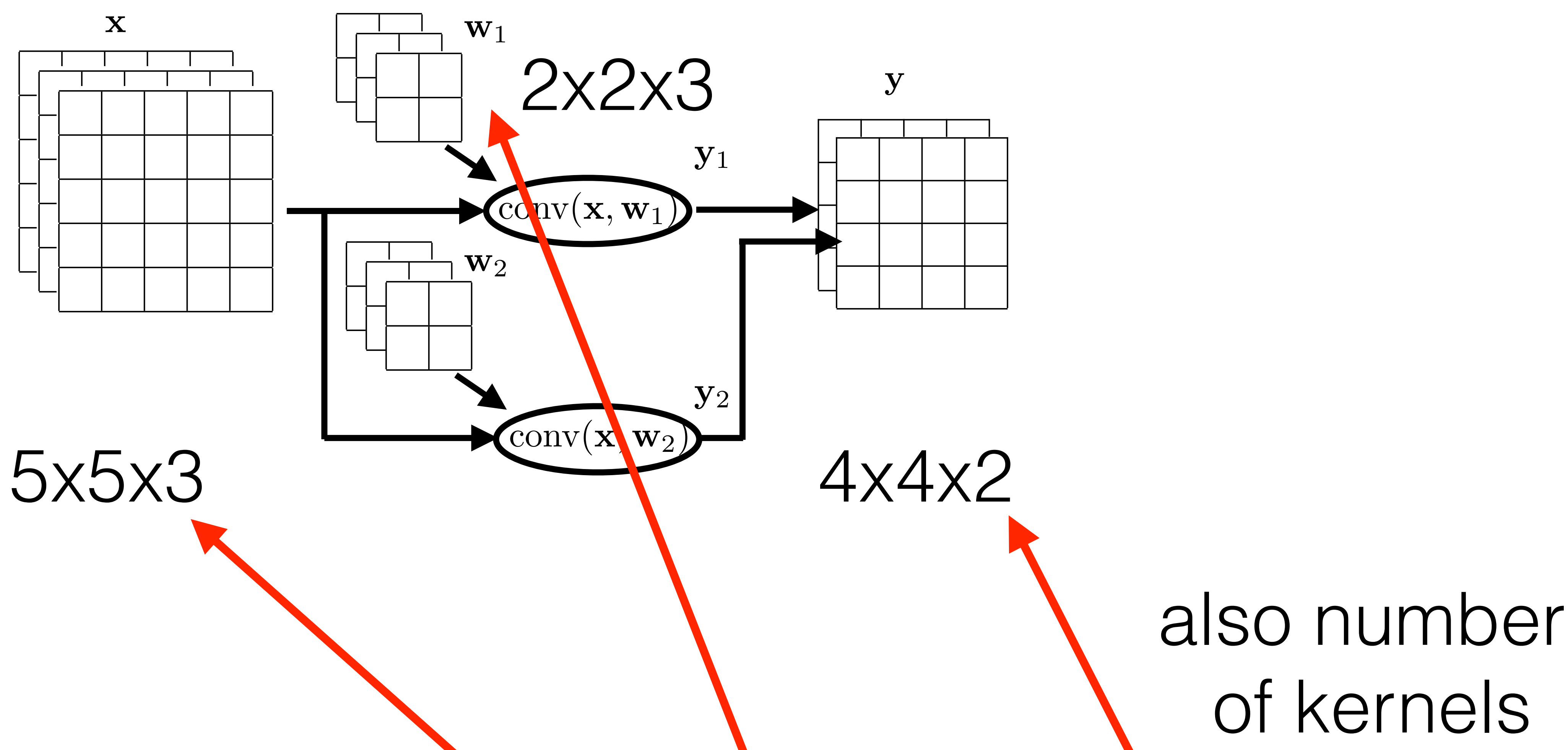
2D convolution forward pass



also number of kernels

```
# initialise
import torch.nn as nn
# define 2D convolutional layer
first_layer = nn.Conv2d(in_channels=3, out_channels=3,
                        kernel_size=2, stride=1,
                        padding=1)
```

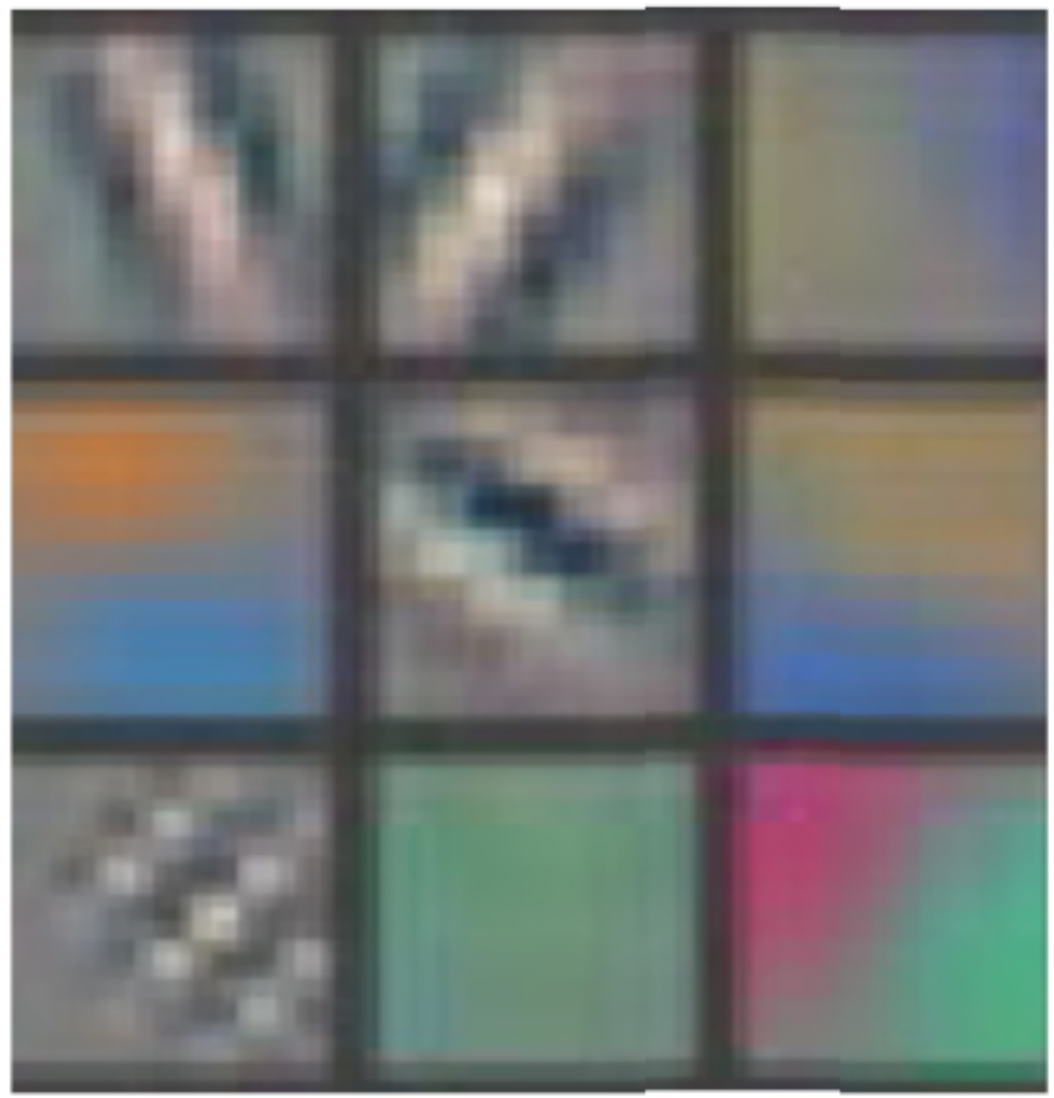
2D convolution forward pass




```
# initialise
import torch.nn as nn
# define 2D convolutional layer
first_layer = nn.Conv2d(in_channels=3, out_channels=2,
                        kernel_size=2, stride=1,
                        padding=1)
```

3. Neurons are sensitive to edges and its orientation

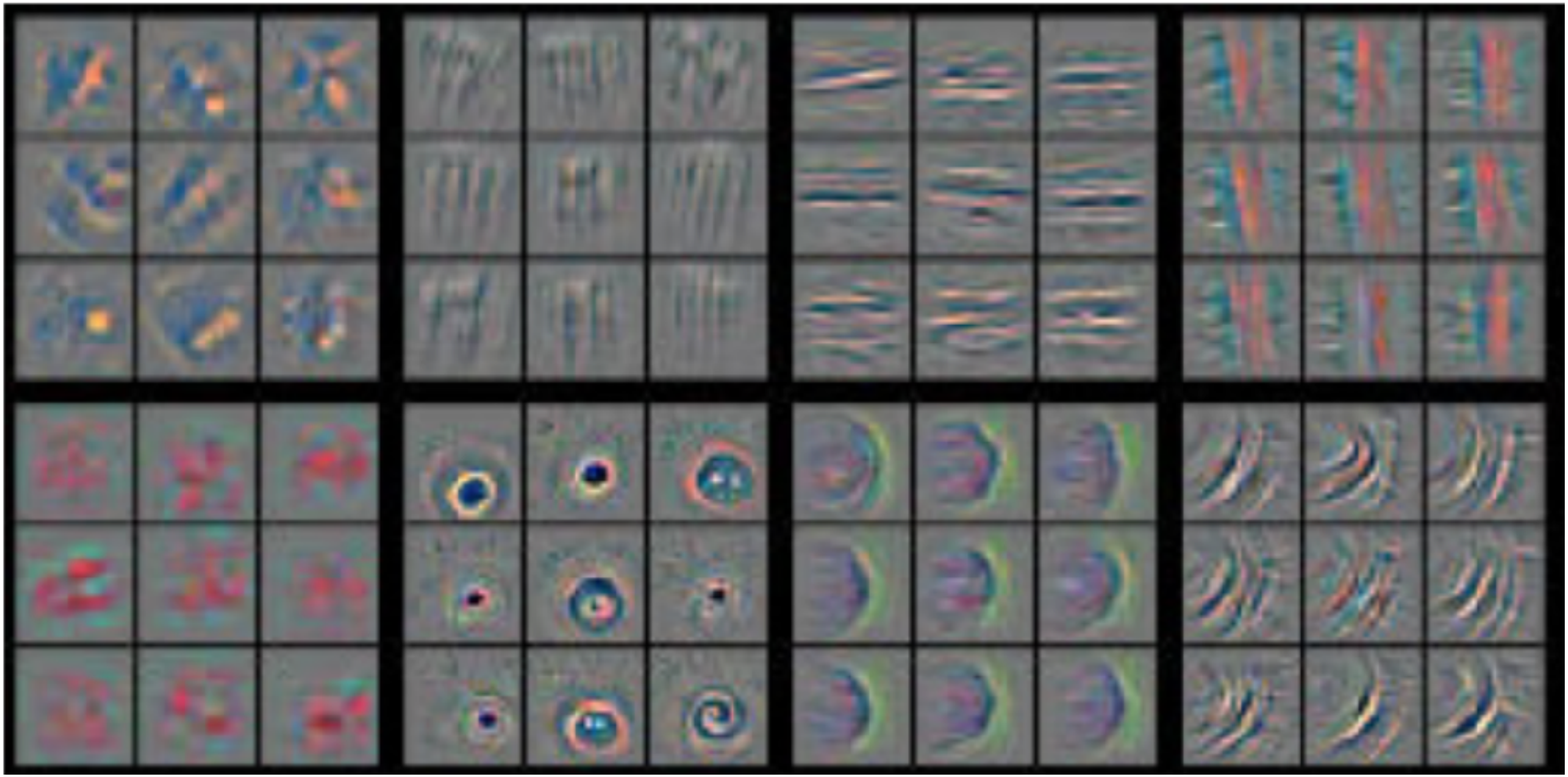
Inputs which maximized output of **layer 1**



[Zeiler and Fergus, ECCV, 2014]

3. Neurons are sensitive to edges and its orientation

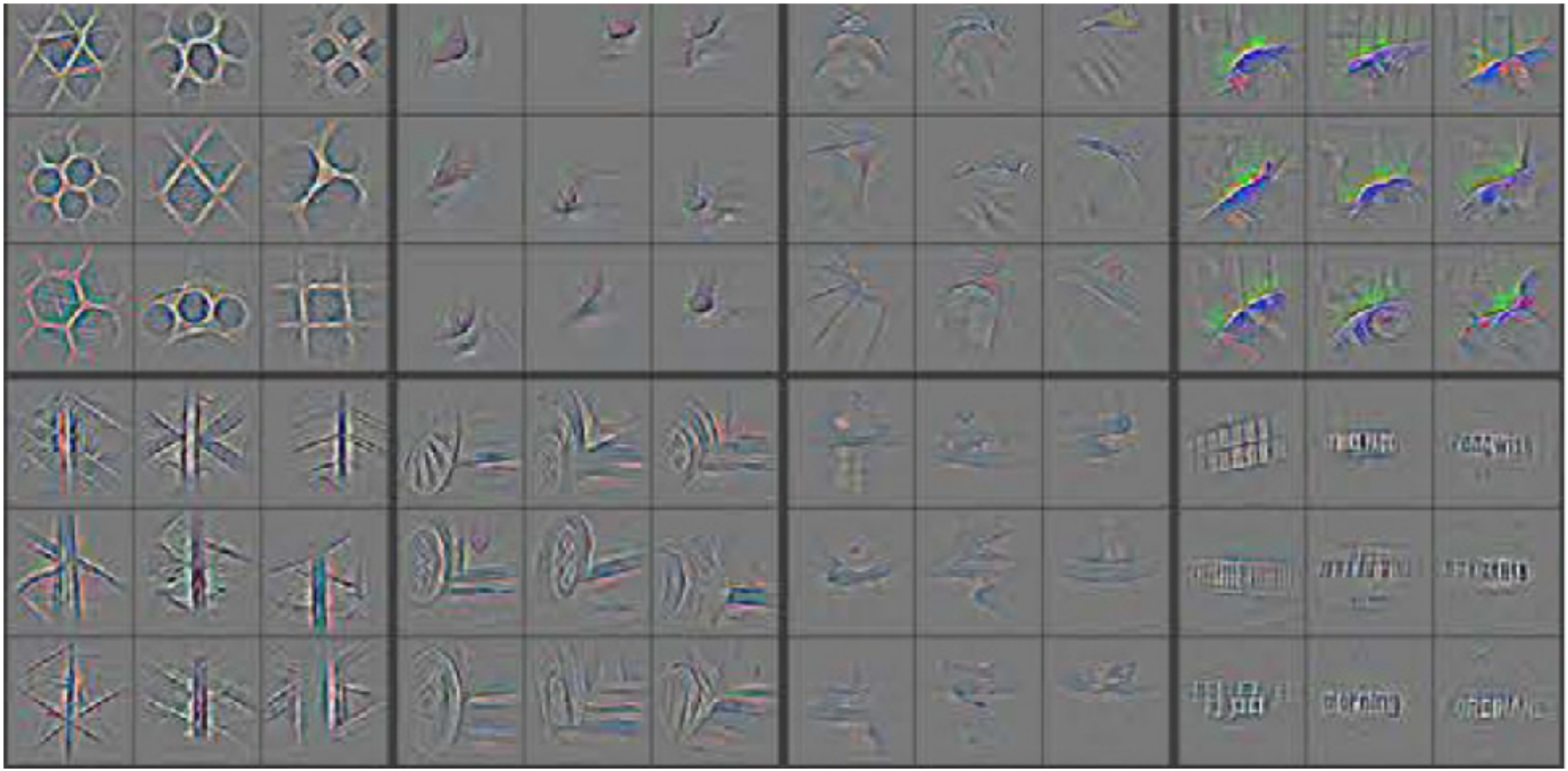
Inputs which maximized output of **layer 2**



[Zeiler and Fergus, ECCV, 2014]

3. Neurons are sensitive to edges and its orientation

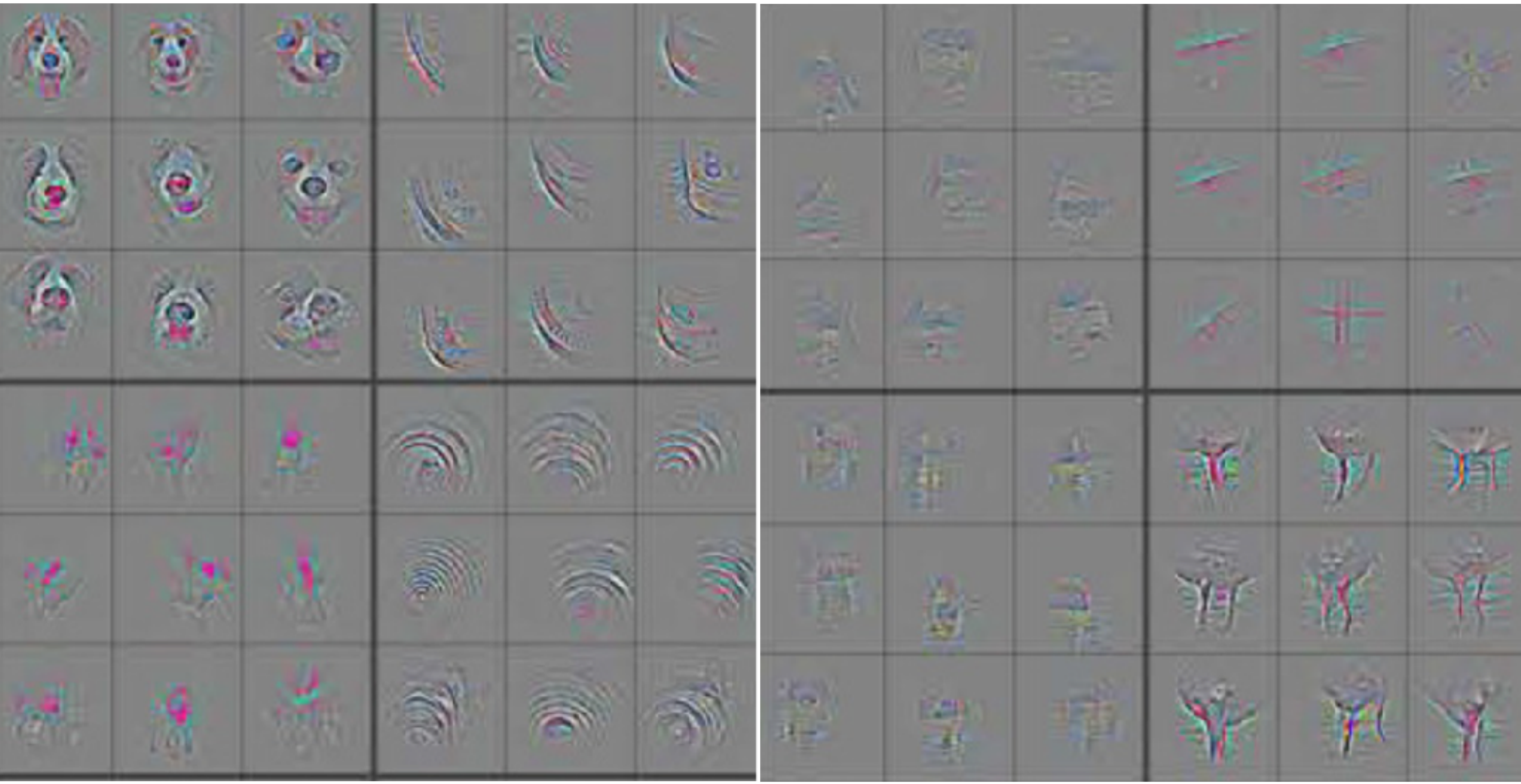
Inputs which maximized output of **layer 3**



[Zeiler and Fergus, ECCV, 2014]

3. Neurons are sensitive to edges and its orientation

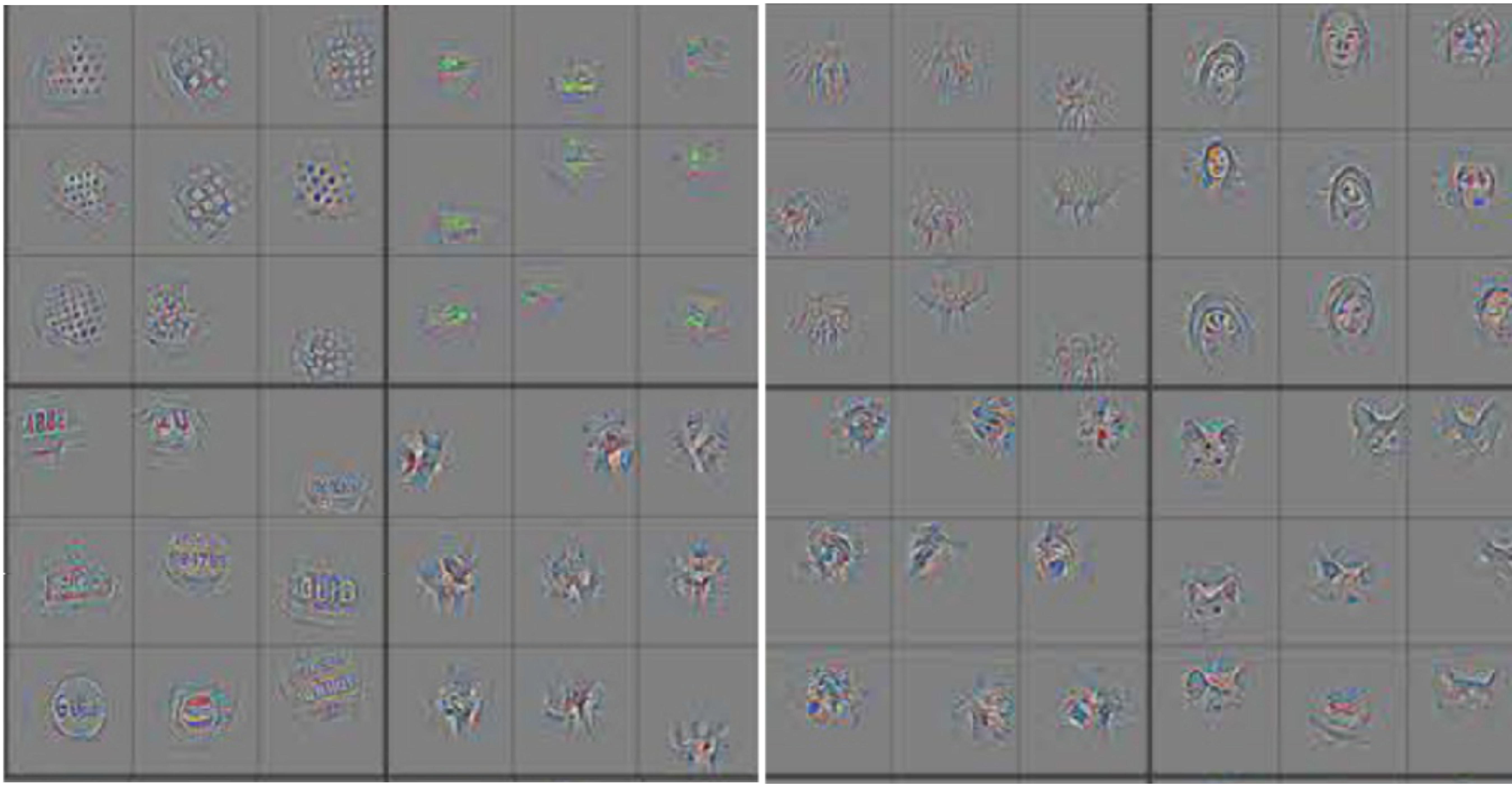
Inputs which maximized output of **layer 4**



[Zeiler and Fergus, ECCV, 2014]

3. Neurons are sensitive to edges and its orientation

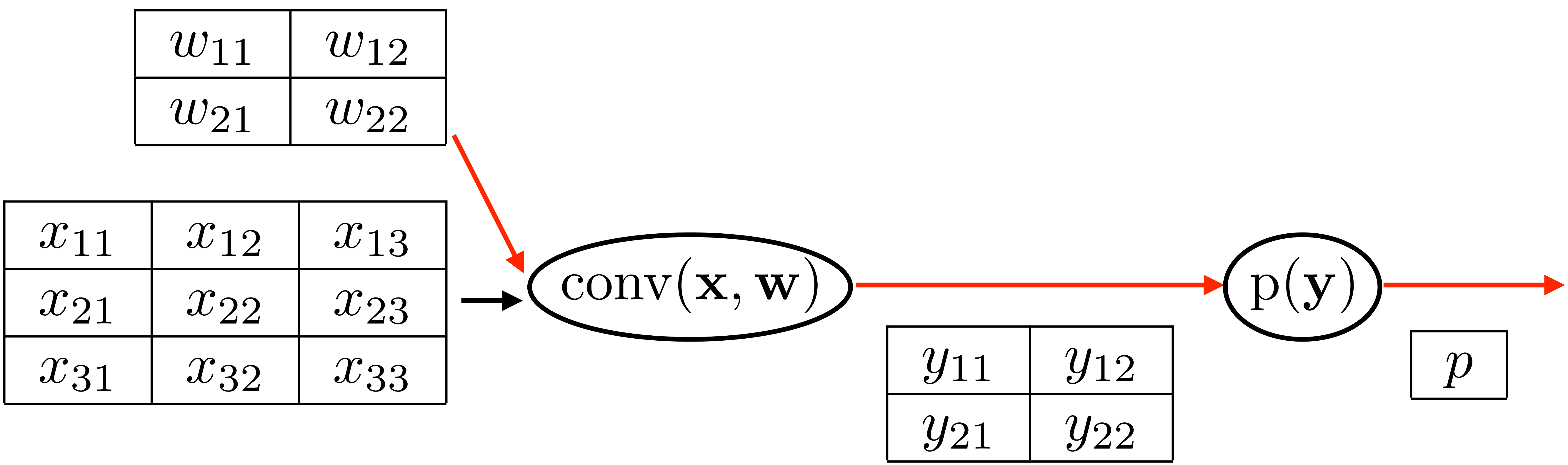
Inputs which maximized output of **layer 5**



[Zeiler and Fergus, ECCV, 2014]

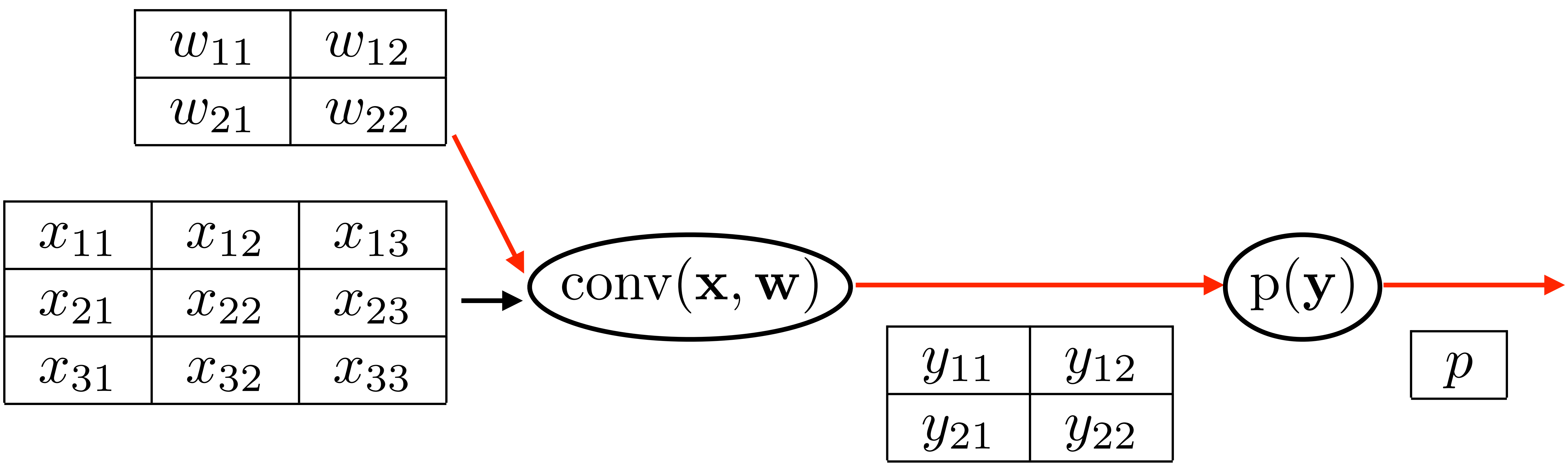
Convolution backward pass

Learning of convolutional neuron => backpropagation



Convolution backward pass

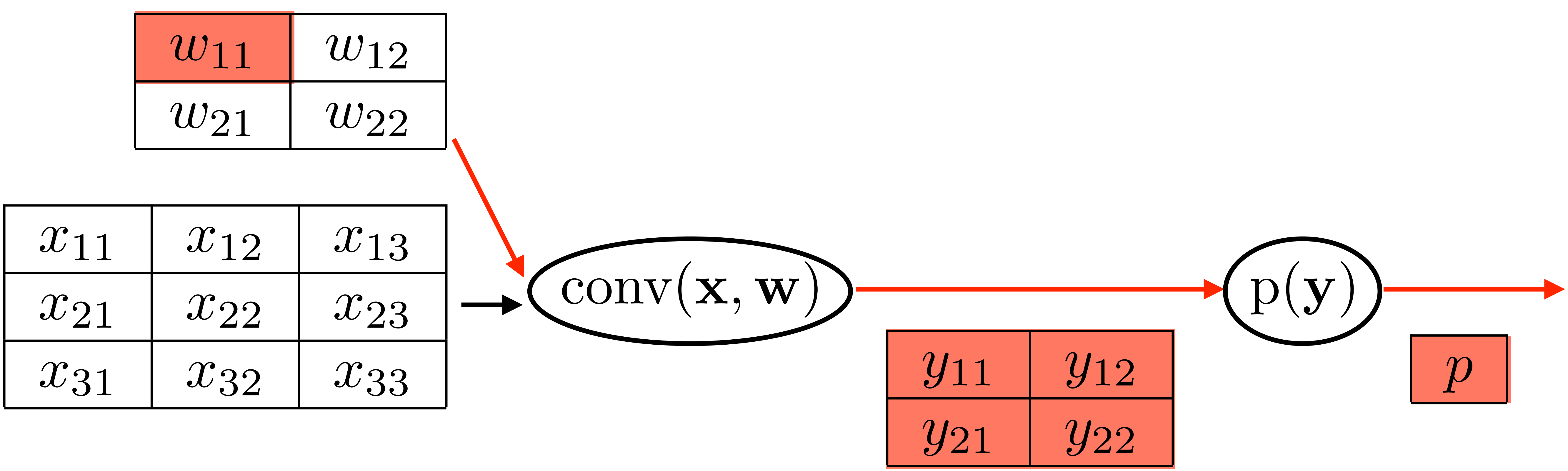
$$\begin{bmatrix} \boxed{\dfrac{\partial p}{\partial w_{11}}} & \dfrac{\partial p}{\partial w_{12}} \\ \dfrac{\partial p}{\partial w_{21}} & \dfrac{\partial p}{\partial w_{22}} \end{bmatrix} = ?$$



Convolution backward pass

$$\begin{array}{|c|c|}\hline \frac{\partial p}{\partial w_{11}} & \frac{\partial p}{\partial w_{12}} \\ \hline \frac{\partial p}{\partial w_{21}} & \frac{\partial p}{\partial w_{22}} \\ \hline \end{array} = ?$$

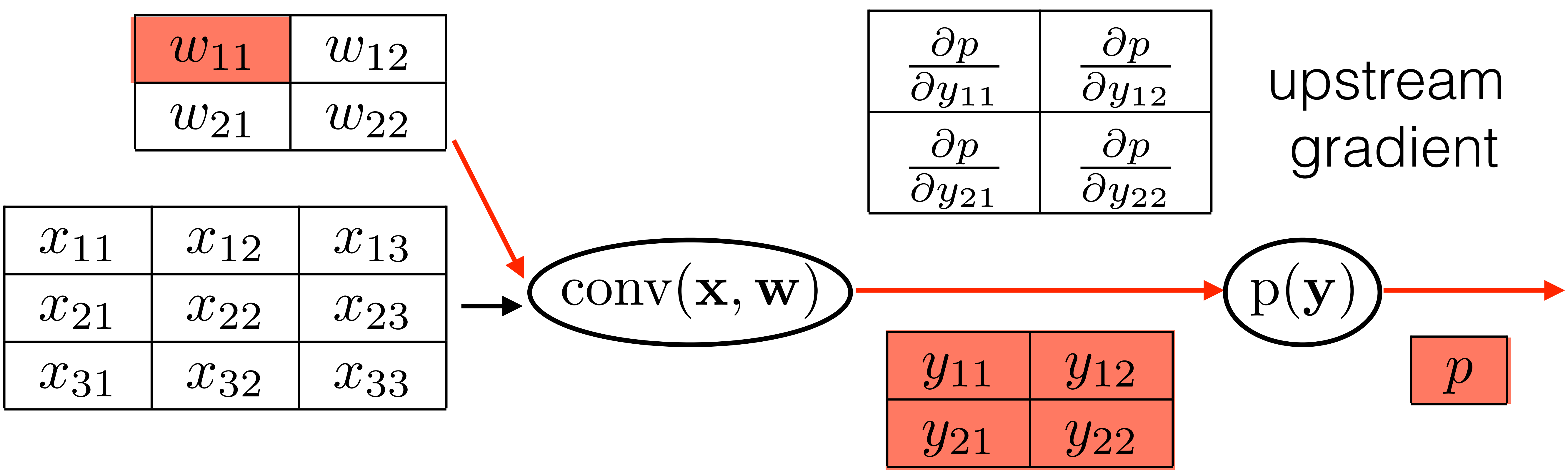
$$p(w_{11}) = \mathrm{p}(y_{11}(w_{11}),\ y_{12}(w_{11}),\ y_{21}(w_{11}),\ y_{22}(w_{11}))$$



Convolution backward pass

$$\frac{\partial p}{\partial w_{11}} = \ \textbf{?}$$

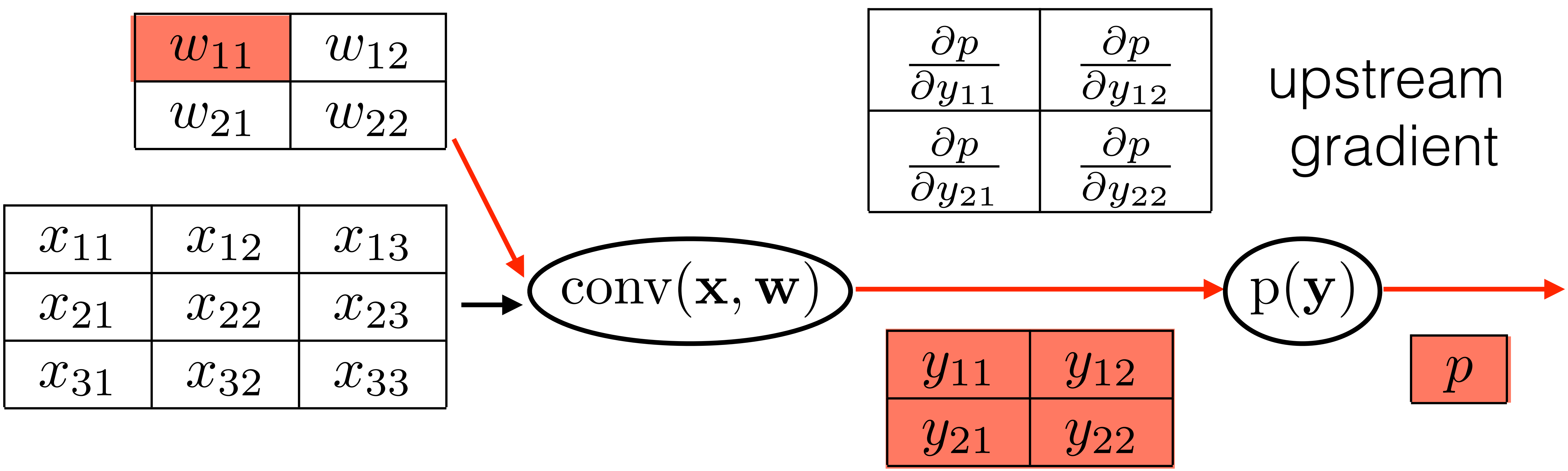
$$p(w_{11}) = \mathrm{p}(y_{11}(w_{11}),\ y_{12}(w_{11}),\ y_{21}(w_{11}),\ y_{22}(w_{11}))$$



Convolution backward pass

$$\frac{\partial p}{\partial w_{11}} = \frac{\partial p}{\partial y_{11}}\frac{\partial y_{11}}{\partial w_{11}} + \frac{\partial p}{\partial y_{12}}\frac{\partial y_{12}}{\partial w_{11}} + \frac{\partial p}{\partial y_{21}}\frac{\partial y_{21}}{\partial w_{11}} + \frac{\partial p}{\partial y_{22}}\frac{\partial y_{22}}{\partial w_{11}}$$

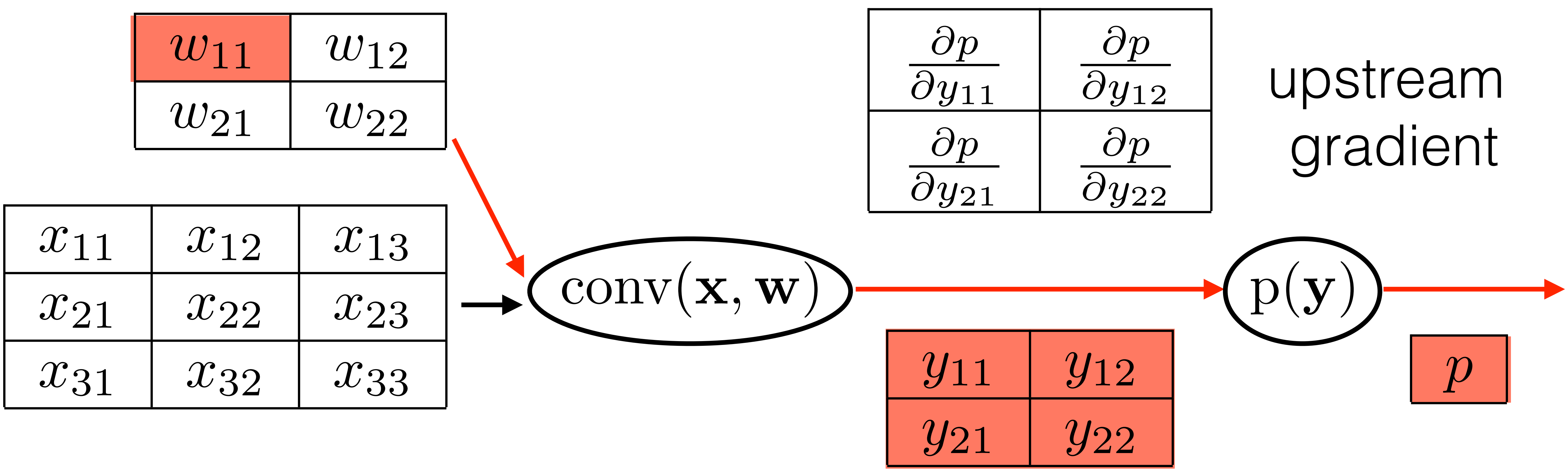
$$p(w_{11}) = \mathrm{p}(y_{11}(w_{11}),\ y_{12}(w_{11}),\ y_{21}(w_{11}),\ y_{22}(w_{11}))$$



Convolution backward pass

$$\frac{\partial p}{\partial w_{11}} = \frac{\partial p}{\partial y_{11}} \boxed{\frac{\partial y_{11}}{\partial w_{11}}} + \frac{\partial p}{\partial y_{12}} \frac{\partial y_{12}}{\partial w_{11}} + \frac{\partial p}{\partial y_{21}} \frac{\partial y_{21}}{\partial w_{11}} + \frac{\partial p}{\partial y_{22}} \frac{\partial y_{22}}{\partial w_{11}}$$

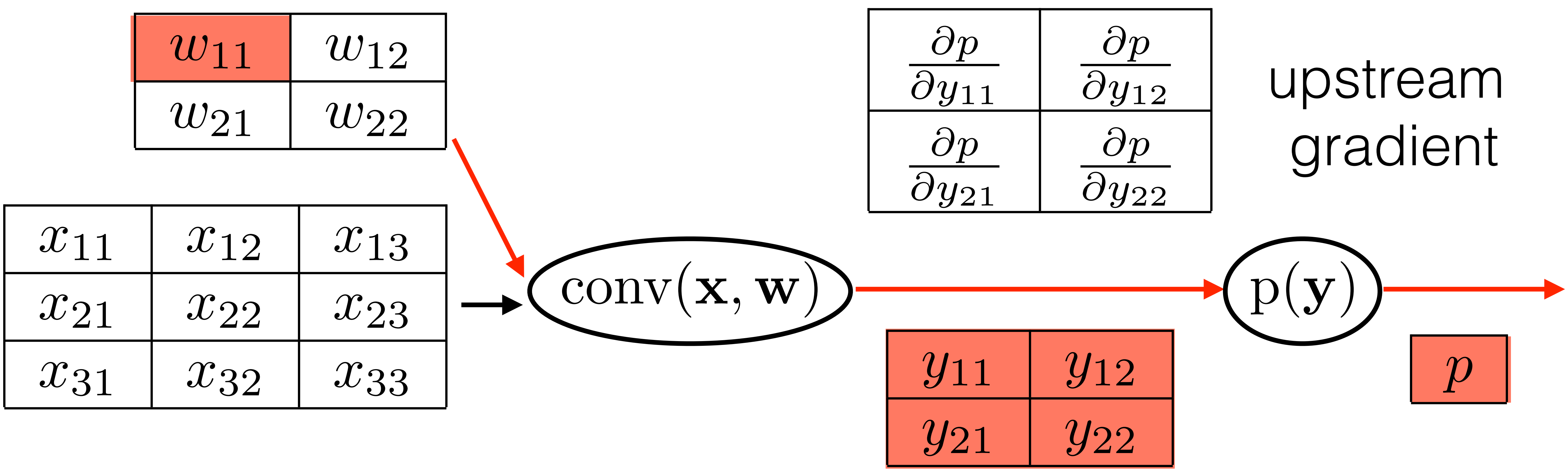
$$\frac{\partial y_{11}}{\partial w_{11}} = \textbf{?}$$



Convolution backward pass

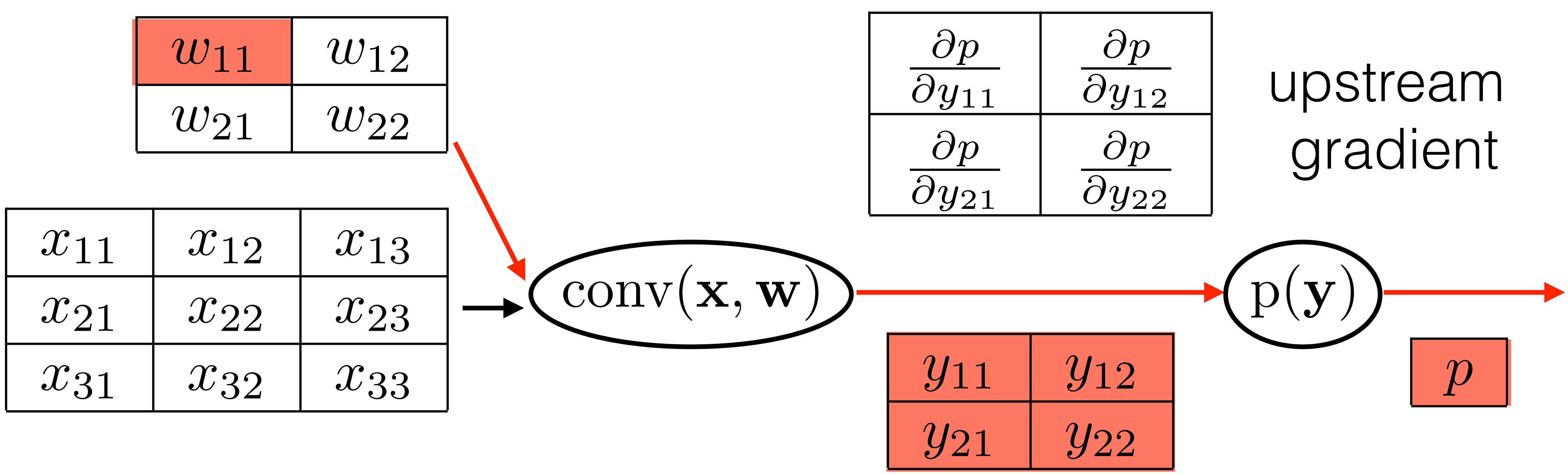
$$\frac{\partial p}{\partial w_{11}} = \frac{\partial p}{\partial y_{11}} \frac{\partial y_{11}}{\partial w_{11}} + \frac{\partial p}{\partial y_{12}} \frac{\partial y_{12}}{\partial w_{11}} + \frac{\partial p}{\partial y_{21}} \frac{\partial y_{21}}{\partial w_{11}} + \frac{\partial p}{\partial y_{22}} \frac{\partial y_{22}}{\partial w_{11}}$$

$$\frac{\partial y_{11}}{\partial w_{11}} = \frac{\partial (w_{11}x_{11} + w_{12}x_{12} + w_{21}x_{21} + w_{22}x_{22})}{\partial w_{11}} = x_{11}$$



Convolution backward pass

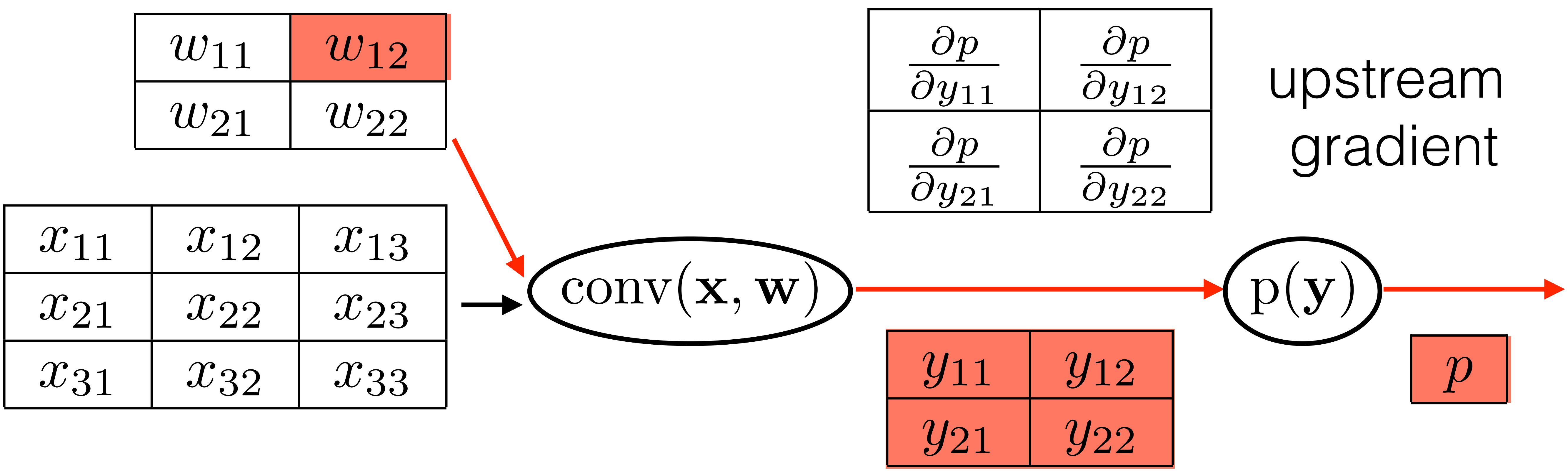
$$\frac{\partial p}{\partial w_{11}} = \frac{\partial p}{\partial y_{11}} x_{11} + \frac{\partial p}{\partial y_{12}} x_{12} + \frac{\partial p}{\partial y_{21}} x_{21} + \frac{\partial p}{\partial y_{22}} x_{22}$$



Convolution backward pass

$$\frac{\partial p}{\partial w_{11}} = \frac{\partial p}{\partial y_{11}} x_{11} + \frac{\partial p}{\partial y_{12}} x_{12} + \frac{\partial p}{\partial y_{21}} x_{21} + \frac{\partial p}{\partial y_{22}} x_{22}$$

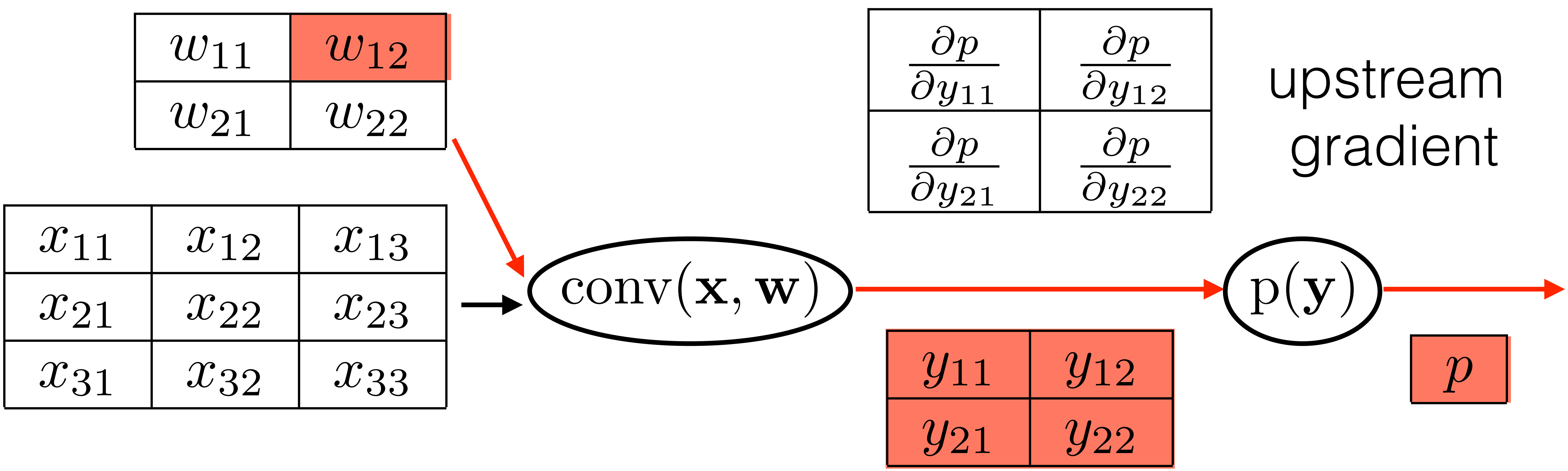
$$\frac{\partial p}{\partial w_{12}} = \textbf{?}$$



Convolution backward pass

$$\frac{\partial p}{\partial w_{11}} = \frac{\partial p}{\partial y_{11}} x_{11} + \frac{\partial p}{\partial y_{12}} x_{12} + \frac{\partial p}{\partial y_{21}} x_{21} + \frac{\partial p}{\partial y_{22}} x_{22}$$

$$\frac{\partial p}{\partial w_{12}} = \frac{\partial p}{\partial y_{11}} x_{12} + \frac{\partial p}{\partial y_{12}} x_{13} + \frac{\partial p}{\partial y_{21}} x_{22} + \frac{\partial p}{\partial y_{22}} x_{23}$$

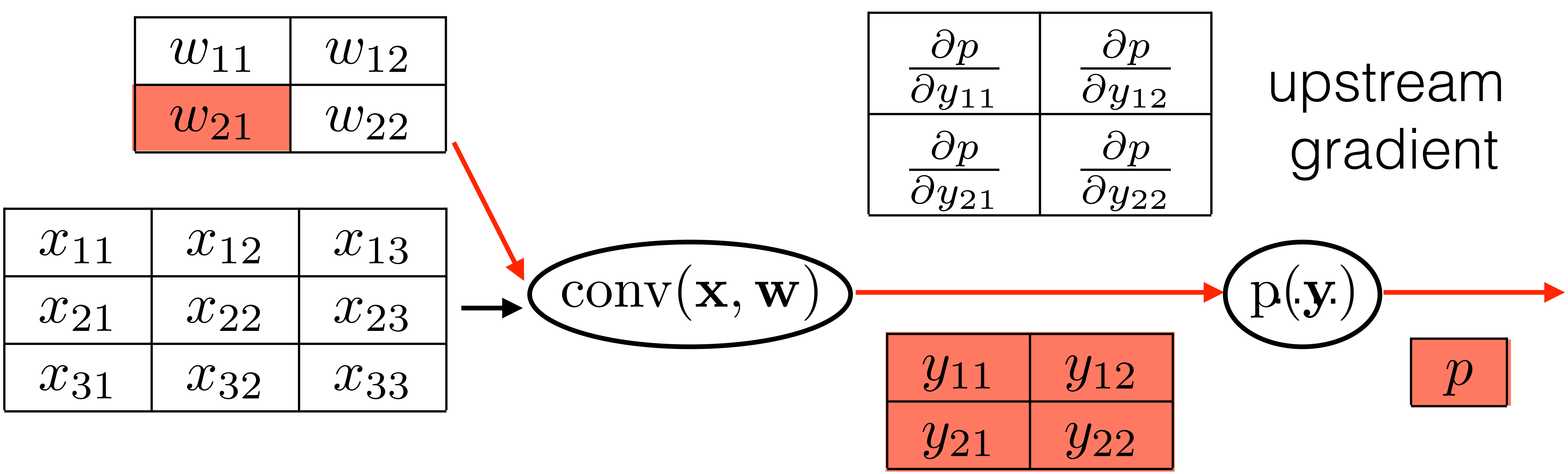


Convolution backward pass

$$\frac{\partial p}{\partial w_{11}} = \frac{\partial p}{\partial y_{11}} x_{11} + \frac{\partial p}{\partial y_{12}} x_{12} + \frac{\partial p}{\partial y_{21}} x_{21} + \frac{\partial p}{\partial y_{22}} x_{22}$$

$$\frac{\partial p}{\partial w_{12}} = \frac{\partial p}{\partial y_{11}} x_{12} + \frac{\partial p}{\partial y_{12}} x_{13} + \frac{\partial p}{\partial y_{21}} x_{22} + \frac{\partial p}{\partial y_{22}} x_{23}$$

$$\frac{\partial p}{\partial w_{21}} = \frac{\partial p}{\partial y_{11}} x_{21} + \frac{\partial p}{\partial y_{12}} x_{22} + \frac{\partial p}{\partial y_{21}} x_{31} + \frac{\partial p}{\partial y_{22}} x_{32}$$



Convolution backward pass

$$\frac{\partial p}{\partial w_{11}} = \frac{\partial p}{\partial y_{11}}x_{11} + \frac{\partial p}{\partial y_{12}}x_{12} + \frac{\partial p}{\partial y_{21}}x_{21} + \frac{\partial p}{\partial y_{22}}x_{22}$$

$$\frac{\partial p}{\partial w_{12}} = \frac{\partial p}{\partial y_{11}}x_{12} + \frac{\partial p}{\partial y_{12}}x_{13} + \frac{\partial p}{\partial y_{21}}x_{22} + \frac{\partial p}{\partial y_{22}}x_{23}$$

$$\frac{\partial p}{\partial w_{21}} = \frac{\partial p}{\partial y_{11}}x_{21} + \frac{\partial p}{\partial y_{12}}x_{22} + \frac{\partial p}{\partial y_{21}}x_{31} + \frac{\partial p}{\partial y_{22}}x_{32}$$

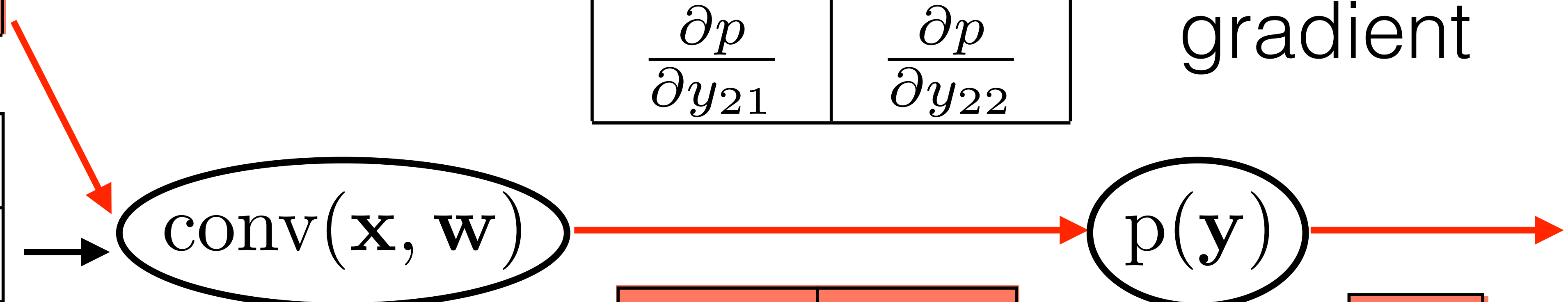
$$\frac{\partial p}{\partial w_{22}} = \frac{\partial p}{\partial y_{11}}x_{22} + \frac{\partial p}{\partial y_{12}}x_{23} + \frac{\partial p}{\partial y_{21}}x_{32} + \frac{\partial p}{\partial y_{22}}x_{33}$$

w_{11}	w_{12}
w_{21}	w_{22}

$\frac{\partial p}{\partial y_{11}}$	$\frac{\partial p}{\partial y_{12}}$
$\frac{\partial p}{\partial y_{21}}$	$\frac{\partial p}{\partial y_{22}}$

upstream gradient

x_{11}	x_{12}	x_{13}
x_{21}	x_{22}	x_{23}
x_{31}	x_{32}	x_{33}



y_{11}	y_{12}
y_{21}	y_{22}

p

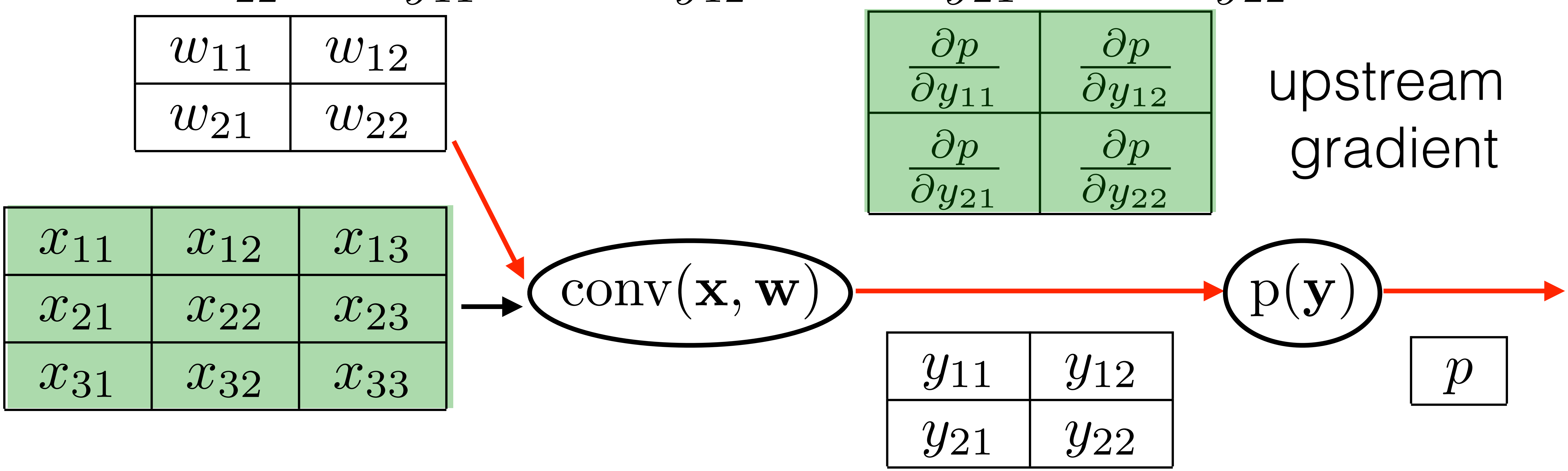
Convolution backward pass

$$\frac{\partial p}{\partial w_{11}} = \frac{\partial p}{\partial y_{11}} x_{11} + \frac{\partial p}{\partial y_{12}} x_{12} + \frac{\partial p}{\partial y_{21}} x_{21} + \frac{\partial p}{\partial y_{22}} x_{22}$$

$$\frac{\partial p}{\partial w_{12}} = \frac{\partial p}{\partial y_{11}} x_{12} + \frac{\partial p}{\partial y_{12}} x_{13} + \frac{\partial p}{\partial y_{21}} x_{22} + \frac{\partial p}{\partial y_{22}} x_{23}$$

$$\frac{\partial p}{\partial w_{21}} = \frac{\partial p}{\partial y_{11}} x_{21} + \frac{\partial p}{\partial y_{12}} x_{22} + \frac{\partial p}{\partial y_{21}} x_{31} + \frac{\partial p}{\partial y_{22}} x_{32}$$

$$\frac{\partial p}{\partial w_{22}} = \frac{\partial p}{\partial y_{11}} x_{22} + \frac{\partial p}{\partial y_{12}} x_{23} + \frac{\partial p}{\partial y_{21}} x_{32} + \frac{\partial p}{\partial y_{22}} x_{33}$$



Convolution backward pass

$$\frac{\partial p}{\partial w_{11}} = \frac{\partial p}{\partial y_{11}}x_{11} + \frac{\partial p}{\partial y_{12}}x_{12} + \frac{\partial p}{\partial y_{21}}x_{21} + \frac{\partial p}{\partial y_{22}}x_{22}$$

$$\frac{\partial p}{\partial w_{12}} = \frac{\partial p}{\partial y_{11}}x_{12} + \frac{\partial p}{\partial y_{12}}x_{13} + \frac{\partial p}{\partial y_{21}}x_{22} + \frac{\partial p}{\partial y_{22}}x_{23}$$

$$\frac{\partial p}{\partial w_{21}} = \frac{\partial p}{\partial y_{11}}x_{21} + \frac{\partial p}{\partial y_{12}}x_{22} + \frac{\partial p}{\partial y_{21}}x_{31} + \frac{\partial p}{\partial y_{22}}x_{32}$$

$$\frac{\partial p}{\partial w_{22}} = \frac{\partial p}{\partial y_{11}}x_{22} + \frac{\partial p}{\partial y_{12}}x_{23} + \frac{\partial p}{\partial y_{21}}x_{32} + \frac{\partial p}{\partial y_{22}}x_{33}$$

$$\begin{array}{|c|c|}\hline \frac{\partial p}{\partial w_{11}} & \frac{\partial p}{\partial w_{12}} \\ \hline \frac{\partial p}{\partial w_{21}} & \frac{\partial p}{\partial w_{22}} \\ \hline \end{array} = \mathrm{conv}\left(\begin{array}{|c|c|c|}\hline x_{11} & x_{12} & x_{13} \\ \hline x_{21} & x_{22} & x_{23} \\ \hline x_{31} & x_{32} & x_{33} \\ \hline \end{array}, \begin{array}{|c|c|}\hline \frac{\partial p}{\partial y_{11}} & \frac{\partial p}{\partial y_{12}} \\ \hline \frac{\partial p}{\partial y_{21}} & \frac{\partial p}{\partial y_{22}} \\ \hline \end{array} \right)$$

Convolution backward pass

$$\frac{\partial p}{\partial w_{11}} = \frac{\partial p}{\partial y_{11}} x_{11} + \frac{\partial p}{\partial y_{12}} x_{12} + \frac{\partial p}{\partial y_{21}} x_{21} + \frac{\partial p}{\partial y_{22}} x_{22}$$

$$\frac{\partial p}{\partial w_{12}} = \frac{\partial p}{\partial y_{11}} x_{12} + \frac{\partial p}{\partial y_{12}} x_{13} + \frac{\partial p}{\partial y_{21}} x_{22} + \frac{\partial p}{\partial y_{22}} x_{23}$$

$$\frac{\partial p}{\partial w_{21}} = \frac{\partial p}{\partial y_{11}} x_{21} + \frac{\partial p}{\partial y_{12}} x_{22} + \frac{\partial p}{\partial y_{21}} x_{31} + \frac{\partial p}{\partial y_{22}} x_{32}$$

$$\frac{\partial p}{\partial w_{22}} = \frac{\partial p}{\partial y_{11}} x_{22} + \frac{\partial p}{\partial y_{12}} x_{23} + \frac{\partial p}{\partial y_{21}} x_{32} + \frac{\partial p}{\partial y_{22}} x_{33}$$

$$\begin{array}{|c|c|}\hline \frac{\partial p}{\partial w_{11}} & \frac{\partial p}{\partial w_{12}} \\ \hline \frac{\partial p}{\partial w_{21}} & \frac{\partial p}{\partial w_{22}} \\ \hline \end{array} = \mathrm{conv}\left(\begin{array}{|c|c|c|}\hline x_{11} & x_{12} & x_{13} \\ \hline x_{21} & x_{22} & x_{23} \\ \hline x_{31} & x_{32} & x_{33} \\ \hline \end{array}, \begin{array}{|c|c|}\hline \frac{\partial p}{\partial y_{11}} & \frac{\partial p}{\partial y_{12}} \\ \hline \frac{\partial p}{\partial y_{21}} & \frac{\partial p}{\partial y_{22}} \\ \hline \end{array} \right)$$

Convolution backward pass

$$\frac{\partial p}{\partial w_{11}} = \frac{\partial p}{\partial y_{11}} x_{11} + \frac{\partial p}{\partial y_{12}} x_{12} + \frac{\partial p}{\partial y_{21}} x_{21} + \frac{\partial p}{\partial y_{22}} x_{22}$$

$$\frac{\partial p}{\partial w_{12}} = \frac{\partial p}{\partial y_{11}} x_{12} + \frac{\partial p}{\partial y_{12}} x_{13} + \frac{\partial p}{\partial y_{21}} x_{22} + \frac{\partial p}{\partial y_{22}} x_{23}$$

$$\frac{\partial p}{\partial w_{21}} = \frac{\partial p}{\partial y_{11}} x_{21} + \frac{\partial p}{\partial y_{12}} x_{22} + \frac{\partial p}{\partial y_{21}} x_{31} + \frac{\partial p}{\partial y_{22}} x_{32}$$

$$\frac{\partial p}{\partial w_{22}} = \frac{\partial p}{\partial y_{11}} x_{22} + \frac{\partial p}{\partial y_{12}} x_{23} + \frac{\partial p}{\partial y_{21}} x_{32} + \frac{\partial p}{\partial y_{22}} x_{33}$$

$$\begin{array}{|c|c|}\hline \frac{\partial p}{\partial w_{11}} & \frac{\partial p}{\partial w_{12}} \\ \hline \frac{\partial p}{\partial w_{21}} & \frac{\partial p}{\partial w_{22}} \\ \hline \end{array} = \mathrm{conv}\left(\begin{array}{|c|c|c|}\hline x_{11} & x_{12} & x_{13} \\ \hline x_{21} & x_{22} & x_{23} \\ \hline x_{31} & x_{32} & x_{33} \\ \hline \end{array}, \begin{array}{|c|c|}\hline \frac{\partial p}{\partial y_{11}} & \frac{\partial p}{\partial y_{12}} \\ \hline \frac{\partial p}{\partial y_{21}} & \frac{\partial p}{\partial y_{22}} \\ \hline \end{array} \right)$$

Convolution backward pass

$$\frac{\partial p}{\partial w_{11}} = \frac{\partial p}{\partial y_{11}} x_{11} + \frac{\partial p}{\partial y_{12}} x_{12} + \frac{\partial p}{\partial y_{21}} x_{21} + \frac{\partial p}{\partial y_{22}} x_{22}$$

$$\frac{\partial p}{\partial w_{12}} = \frac{\partial p}{\partial y_{11}} x_{12} + \frac{\partial p}{\partial y_{12}} x_{13} + \frac{\partial p}{\partial y_{21}} x_{22} + \frac{\partial p}{\partial y_{22}} x_{23}$$

$$\frac{\partial p}{\partial w_{21}} = \frac{\partial p}{\partial y_{11}} x_{21} + \frac{\partial p}{\partial y_{12}} x_{22} + \frac{\partial p}{\partial y_{21}} x_{31} + \frac{\partial p}{\partial y_{22}} x_{32}$$

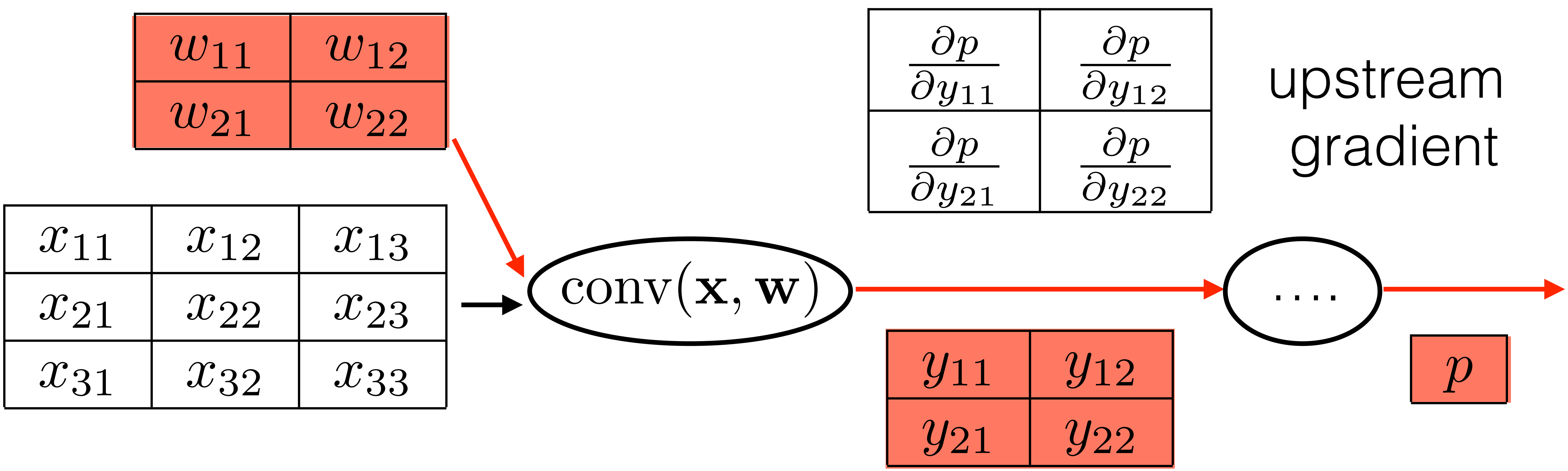
$$\frac{\partial p}{\partial w_{22}} = \frac{\partial p}{\partial y_{11}} x_{22} + \frac{\partial p}{\partial y_{12}} x_{23} + \frac{\partial p}{\partial y_{21}} x_{32} + \frac{\partial p}{\partial y_{22}} x_{33}$$

$$\begin{array}{|c|c|}\hline \frac{\partial p}{\partial w_{11}} & \frac{\partial p}{\partial w_{12}} \\ \hline \frac{\partial p}{\partial w_{21}} & \frac{\partial p}{\partial w_{22}} \\ \hline \end{array} = \text{conv}\left(\begin{array}{|c|c|c|}\hline x_{11} & x_{12} & x_{13} \\ \hline x_{21} & x_{22} & x_{23} \\ \hline x_{31} & x_{32} & x_{33} \\ \hline \end{array}, \begin{array}{|c|c|}\hline \frac{\partial p}{\partial y_{11}} & \frac{\partial p}{\partial y_{12}} \\ \hline \frac{\partial p}{\partial y_{21}} & \frac{\partial p}{\partial y_{22}} \\ \hline \end{array} \right)$$

Convolution backward pass wrt weights

- Backpropagation in convolutional layer wrt weights is: "**convolution of input feature map with upstream gradient**"

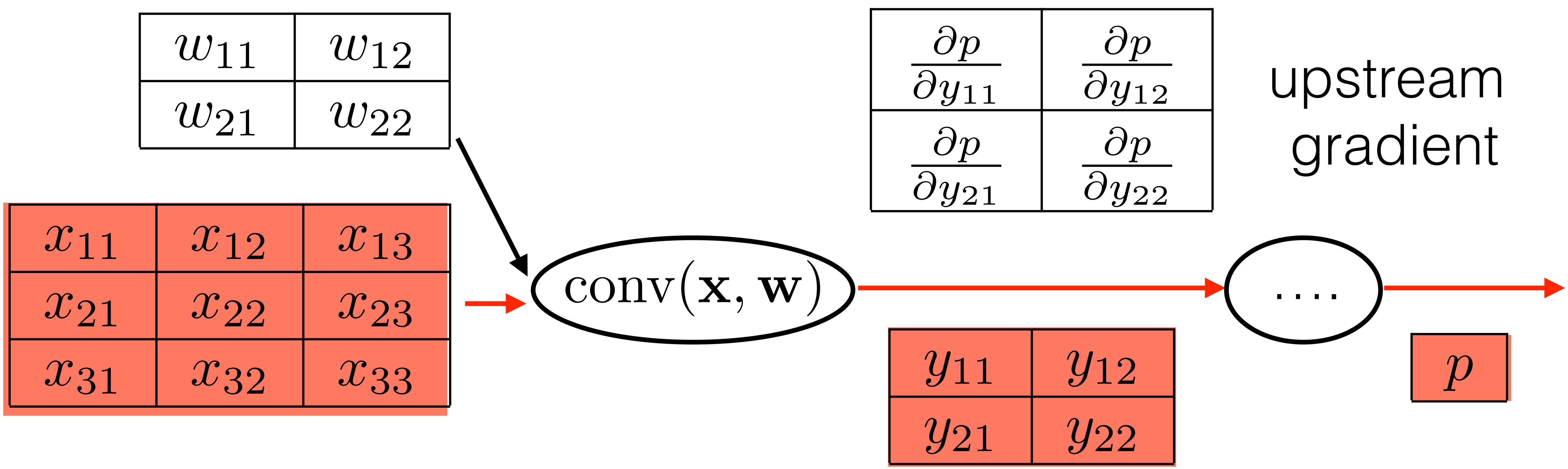
$$\begin{bmatrix} \frac{\partial p}{\partial w_{11}} & \frac{\partial p}{\partial w_{12}} \\ \frac{\partial p}{\partial w_{21}} & \frac{\partial p}{\partial w_{22}} \end{bmatrix} = \mathrm{conv}\left(\begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{bmatrix}, \begin{bmatrix} \frac{\partial p}{\partial y_{11}} & \frac{\partial p}{\partial y_{12}} \\ \frac{\partial p}{\partial y_{21}} & \frac{\partial p}{\partial y_{22}} \end{bmatrix} \right)$$



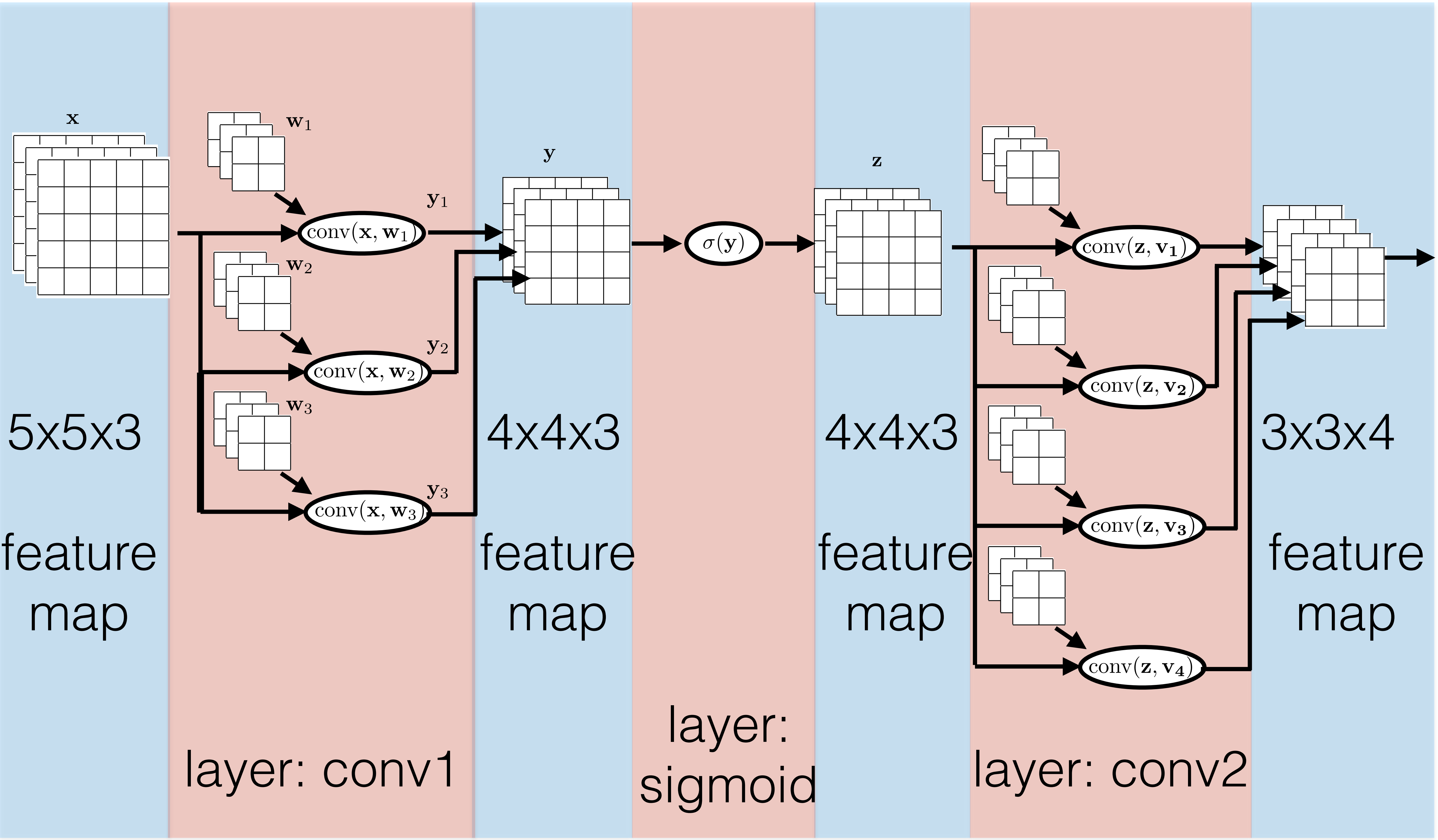
Convolution backward pass wrt input feature map

Backpropagation in convolutional layer is:
"**convolution of padded upstream gradient with mirrored weights**"

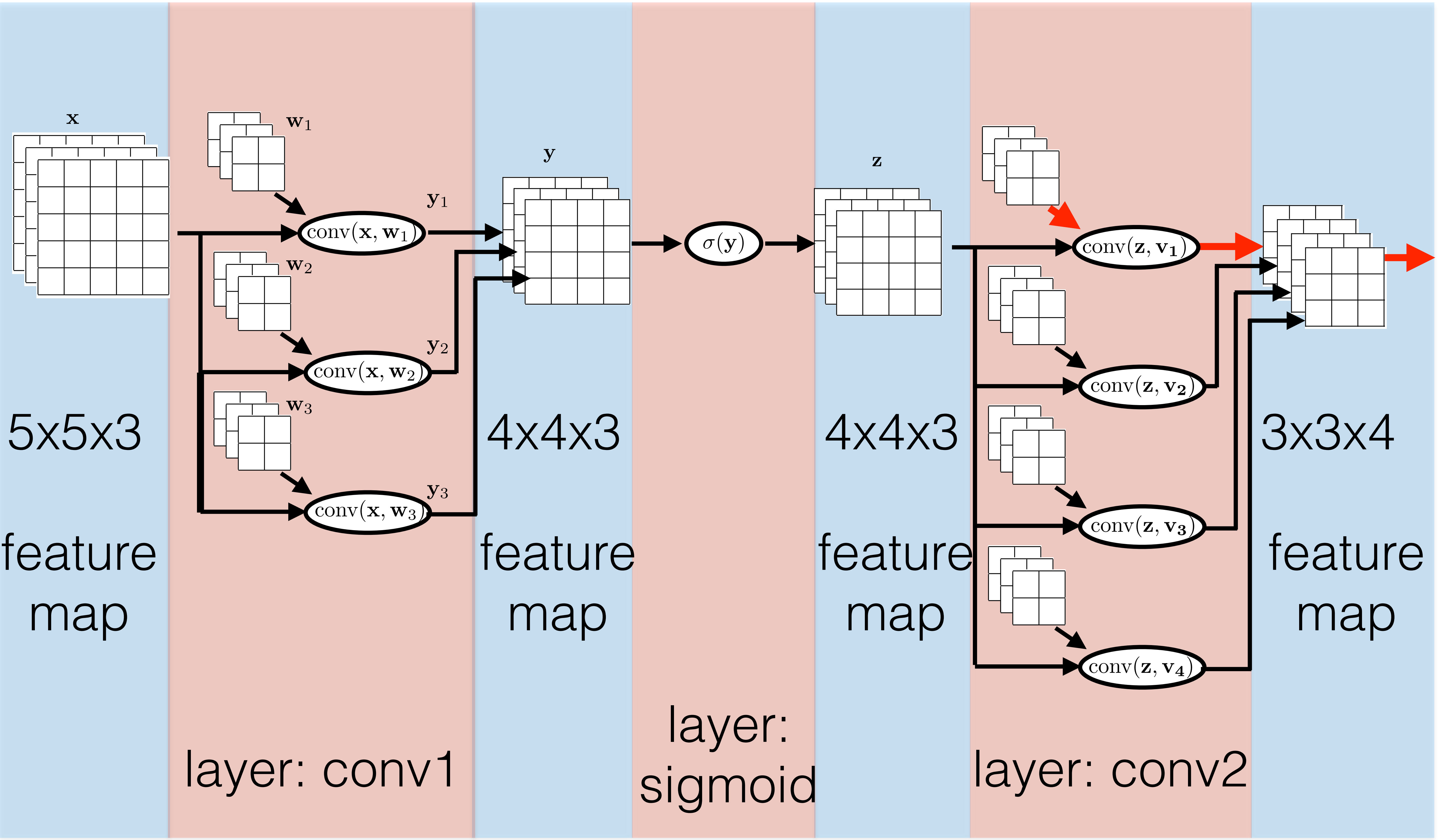
$$\begin{bmatrix} \frac{\partial p}{\partial x_{11}} & \frac{\partial p}{\partial x_{12}} & \frac{\partial p}{\partial x_{13}} \\ \frac{\partial p}{\partial x_{21}} & \frac{\partial p}{\partial x_{22}} & \frac{\partial p}{\partial x_{23}} \\ \frac{\partial p}{\partial x_{31}} & \frac{\partial p}{\partial x_{32}} & \frac{\partial p}{\partial x_{33}} \end{bmatrix} = \mathrm{conv}\left(\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{\partial p}{\partial y_{11}} & \frac{\partial p}{\partial y_{12}} & 0 \\ 0 & \frac{\partial p}{\partial y_{21}} & \frac{\partial p}{\partial y_{22}} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} w_{22} & w_{21} \\ w_{12} & w_{11} \end{bmatrix} \right)$$



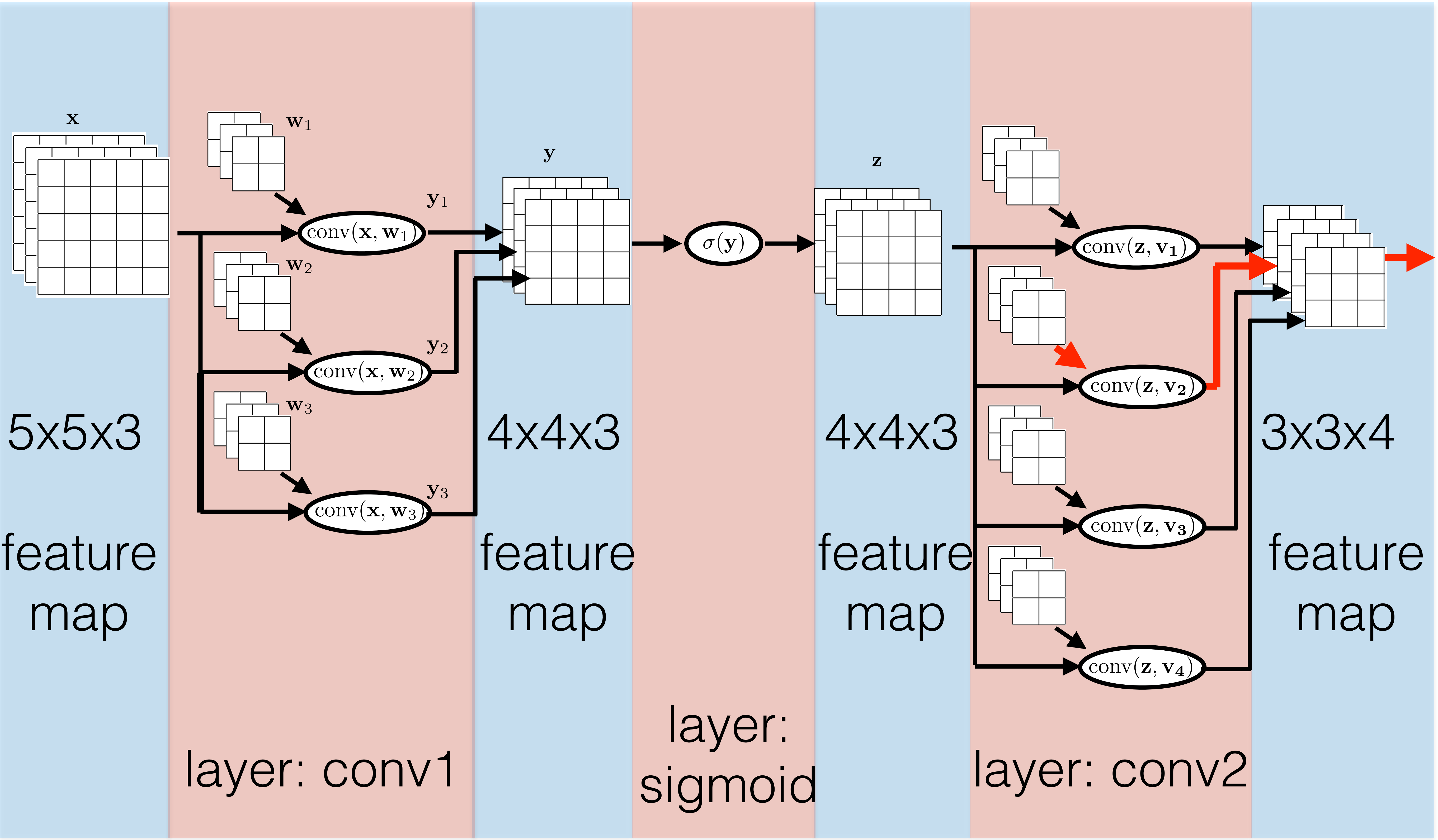
???? Convolutional network backprop ?????



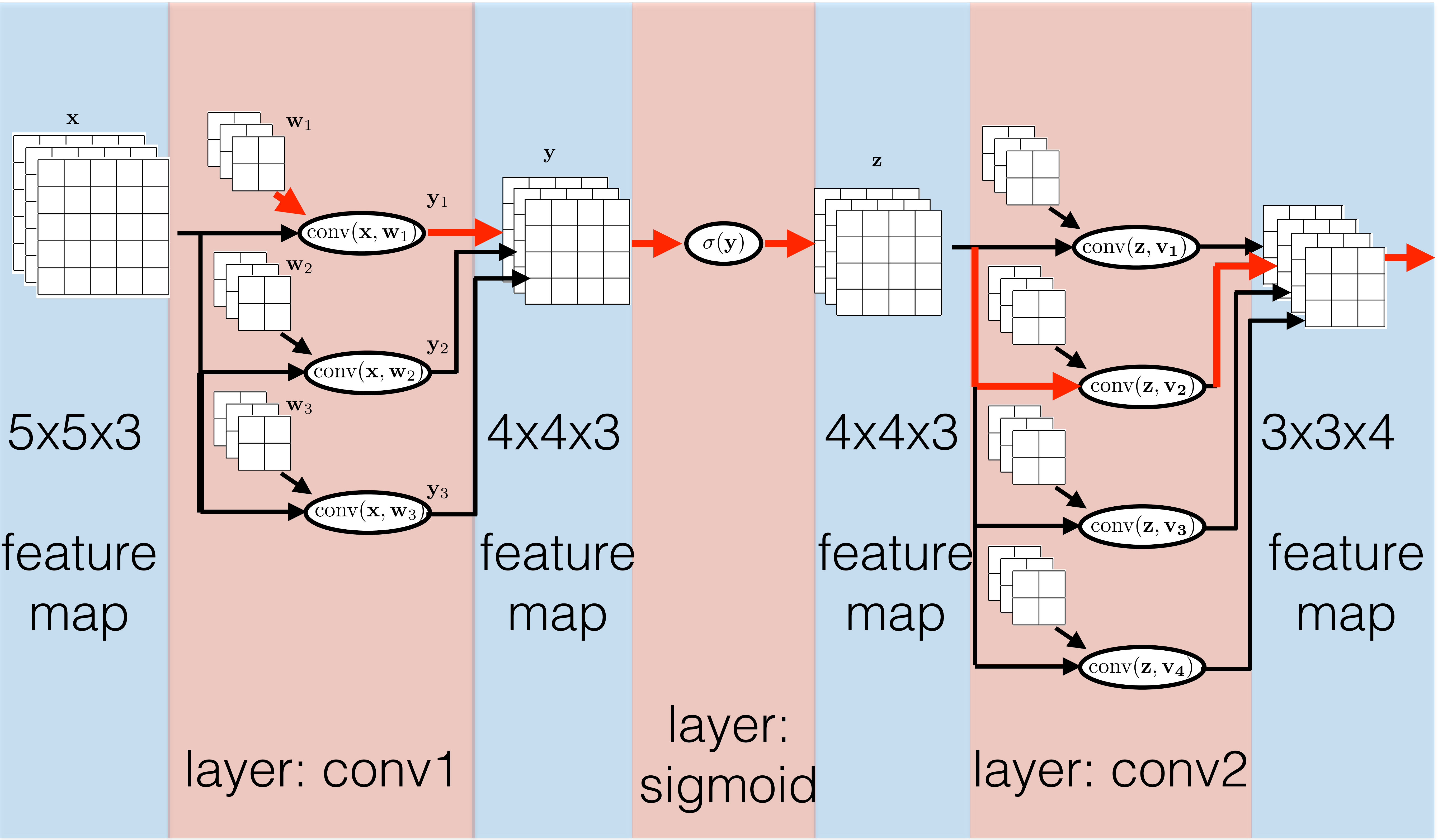
???? Convolutional network backprop ?????



???? Convolutional network backprop ?????



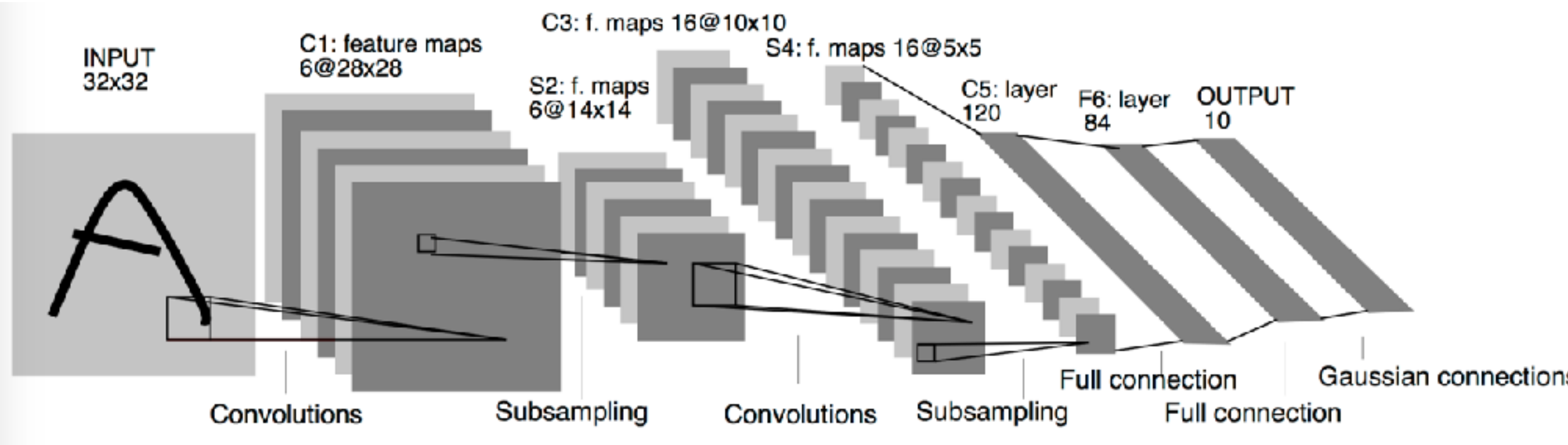
???? Convolutional network backprop ?????



Convolutional net

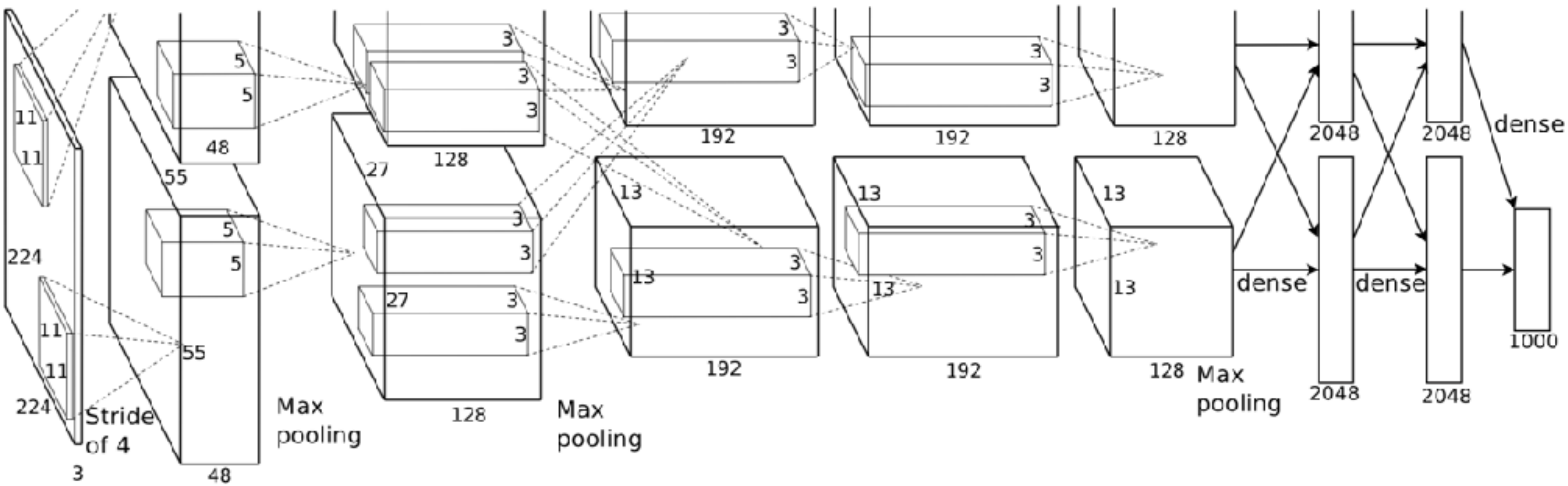
- Convolutional network (ConvNet) is concatenation of convolutional layers
- Backprop in ConvNet is convolution of feature maps or kernels or feature-maps with the upstream gradient.
- Feed-forward and backprop are convolutions => efficient implementation on GPU

LeCun's letter recognition 1998 (over 13k citations !!!)



LeCun et al, Gradient based learning applied to document recognition, IEEE, 1998
http://yann.lecun.com/exdb/publis/pdf/lecun-01a.pdf

AlexNet on ImageNet 2012 (**over 27k citations !!!**)



Alex Krizhevsky et al, Imagenet classification with deep convolutional neural networks, NIPS, 2012
https://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf

IMAGENET

Classification results

http://image-net.org/challenges/LSVRC/2017/index

Steel drum

Output:
Scale
T-shirt
<u>Steel drum</u>
Drumstick
Mud turtle

✔

Output:
Scale
T-shirt
Giant panda
Drumstick
Mud turtle

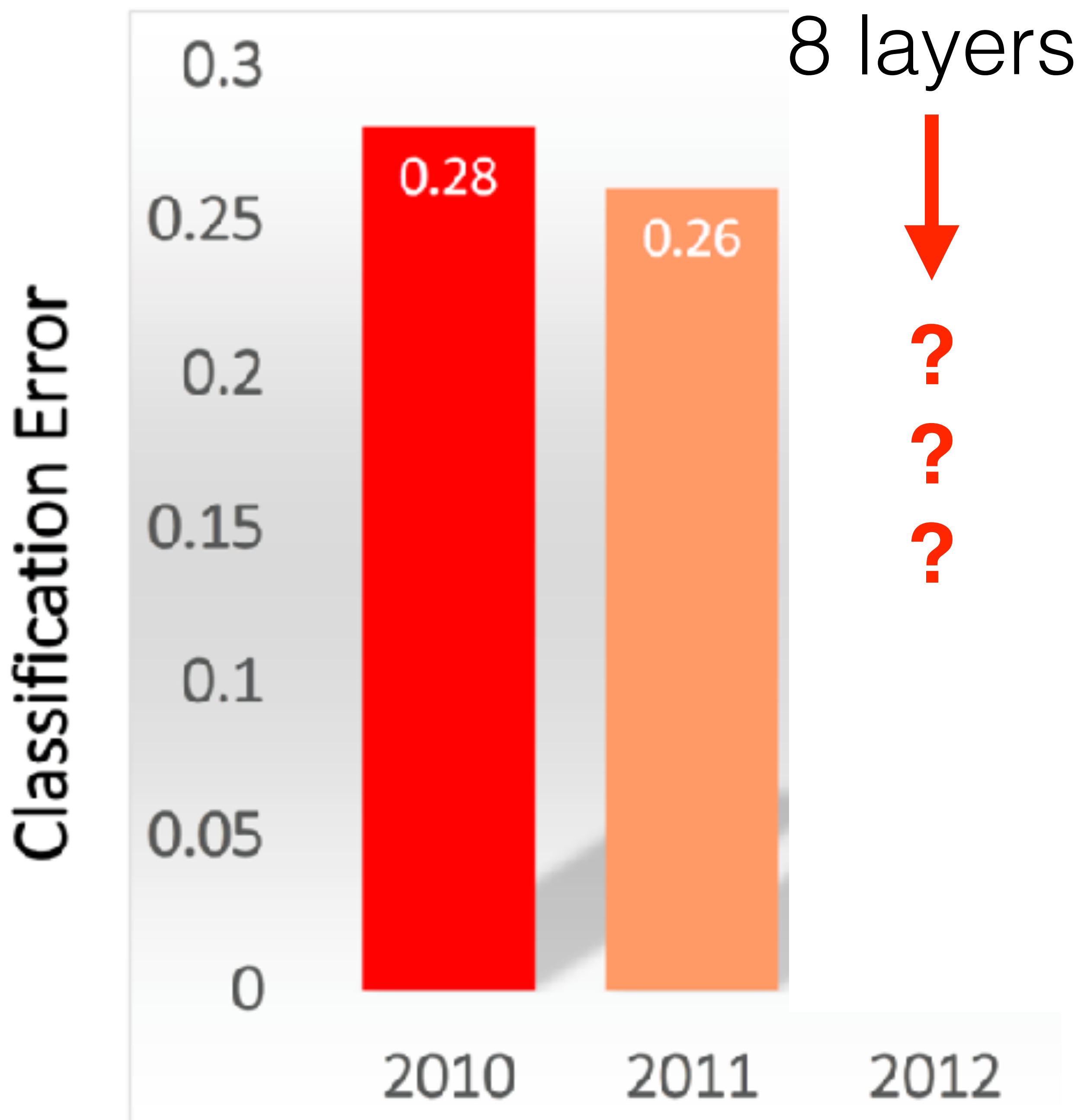
✗

$$\text{Error} = \frac{1}{100{,}000} \sum_{\substack{100{,}000 \\ \text{images}}} 1[\text{incorrect on image } i]$$

IMAGENET

Classification results

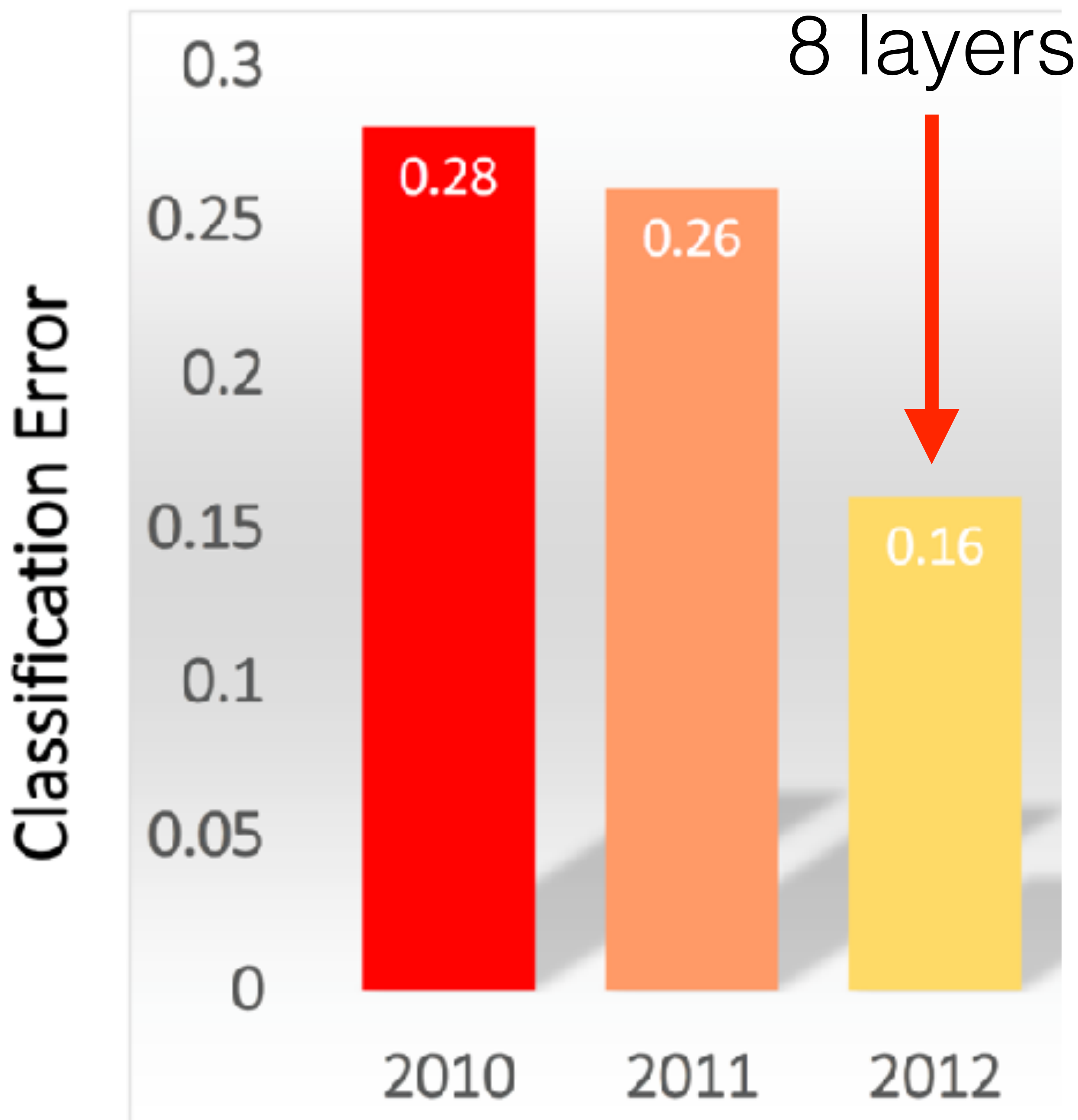
AlexNet
8 layers



IMAGENET

Classification results

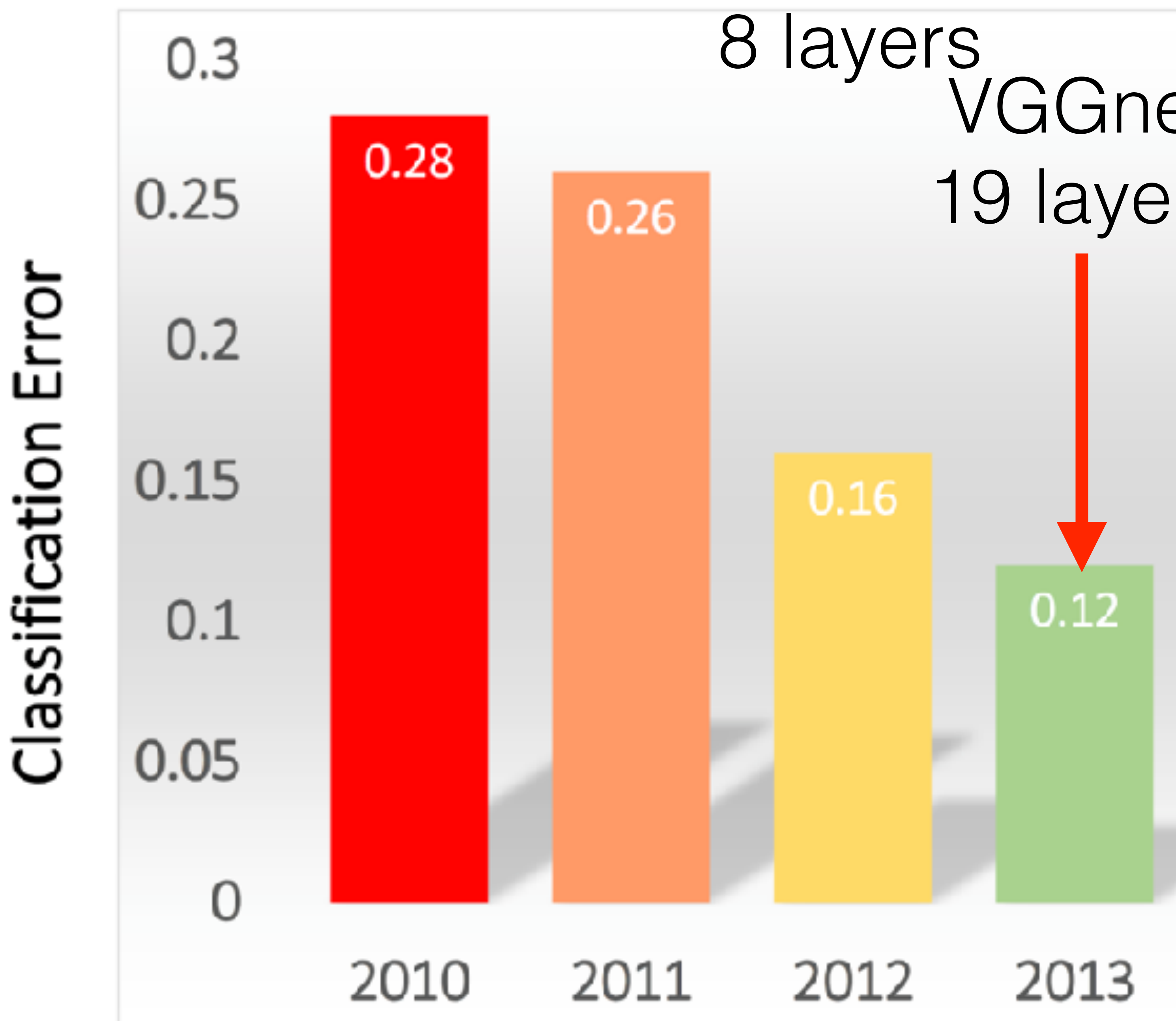
AlexNet
8 layers



Classification results

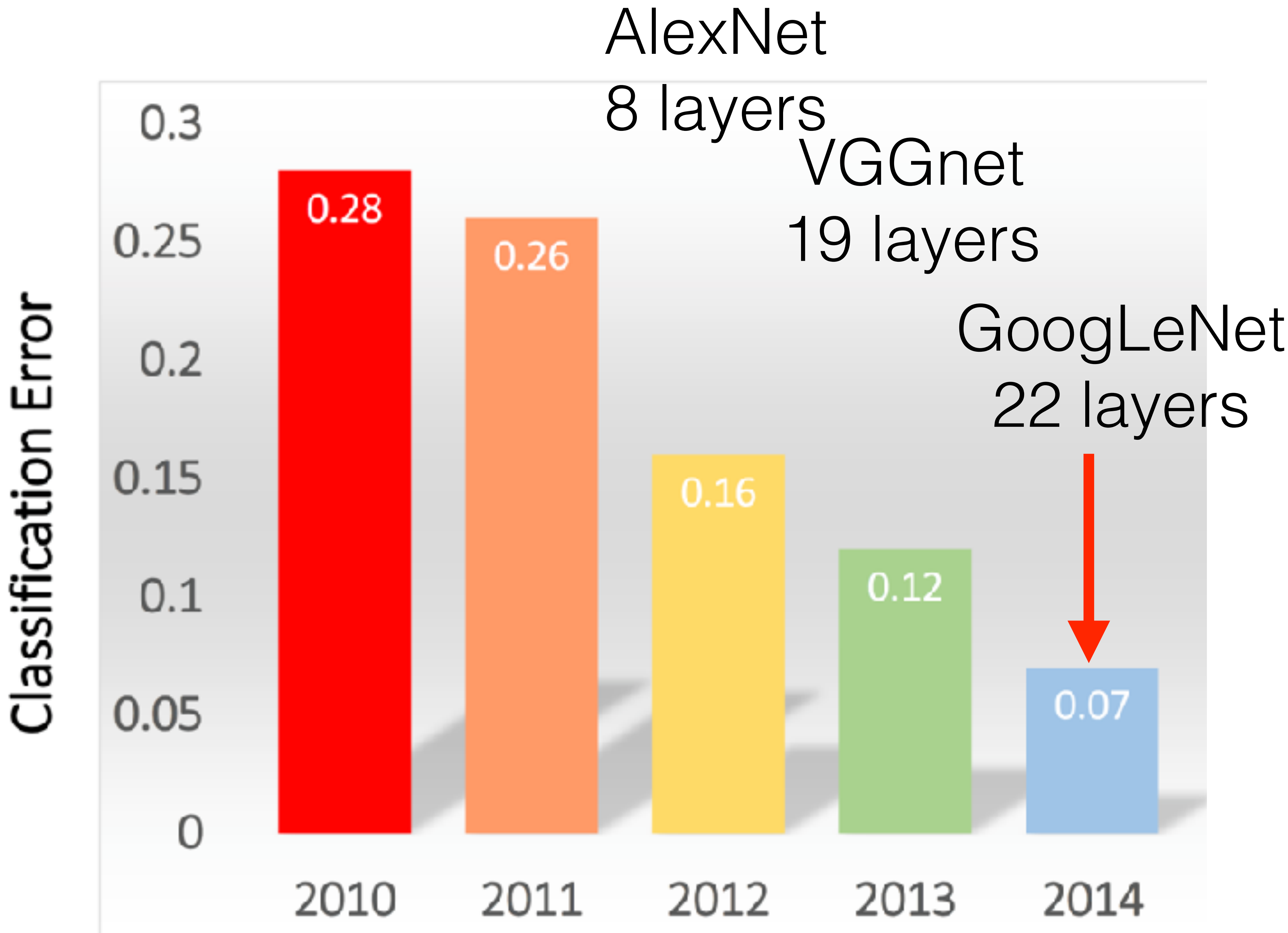
AlexNet
8 layers

VGGnet
19 layers



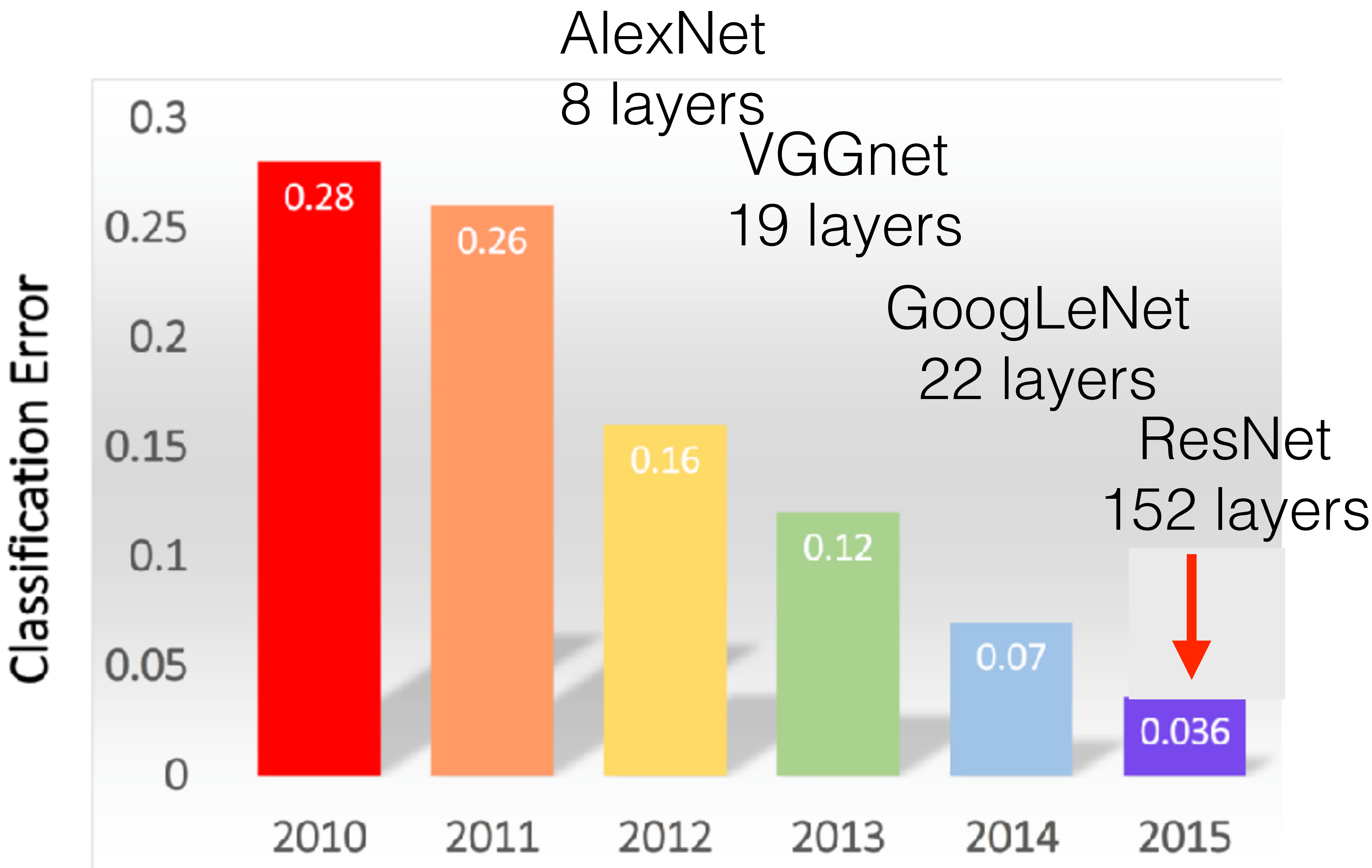
IMAGENET

Classification results



IMAGENET

Classification results



IMAGENET

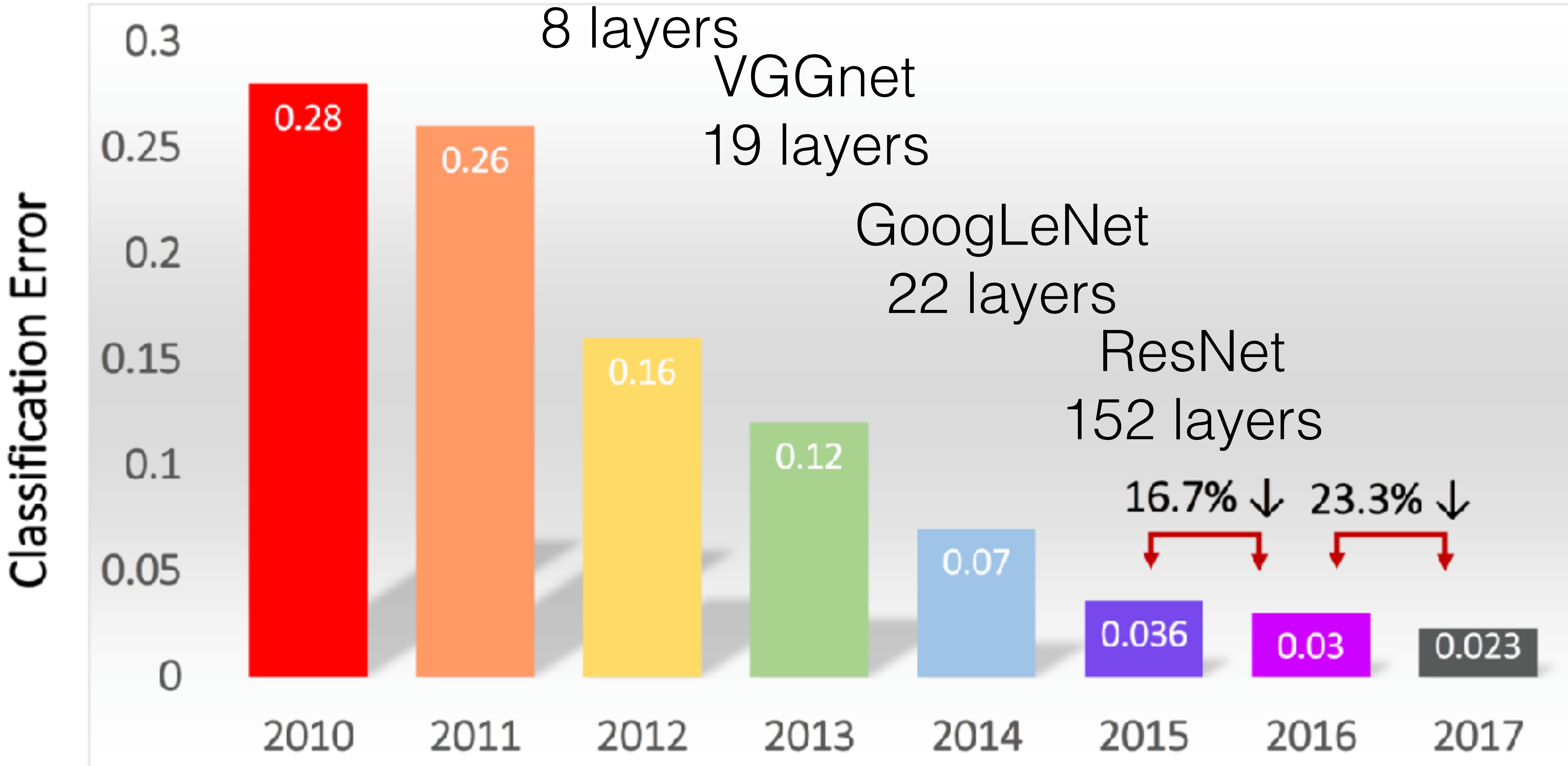
Classification results

AlexNet
8 layers

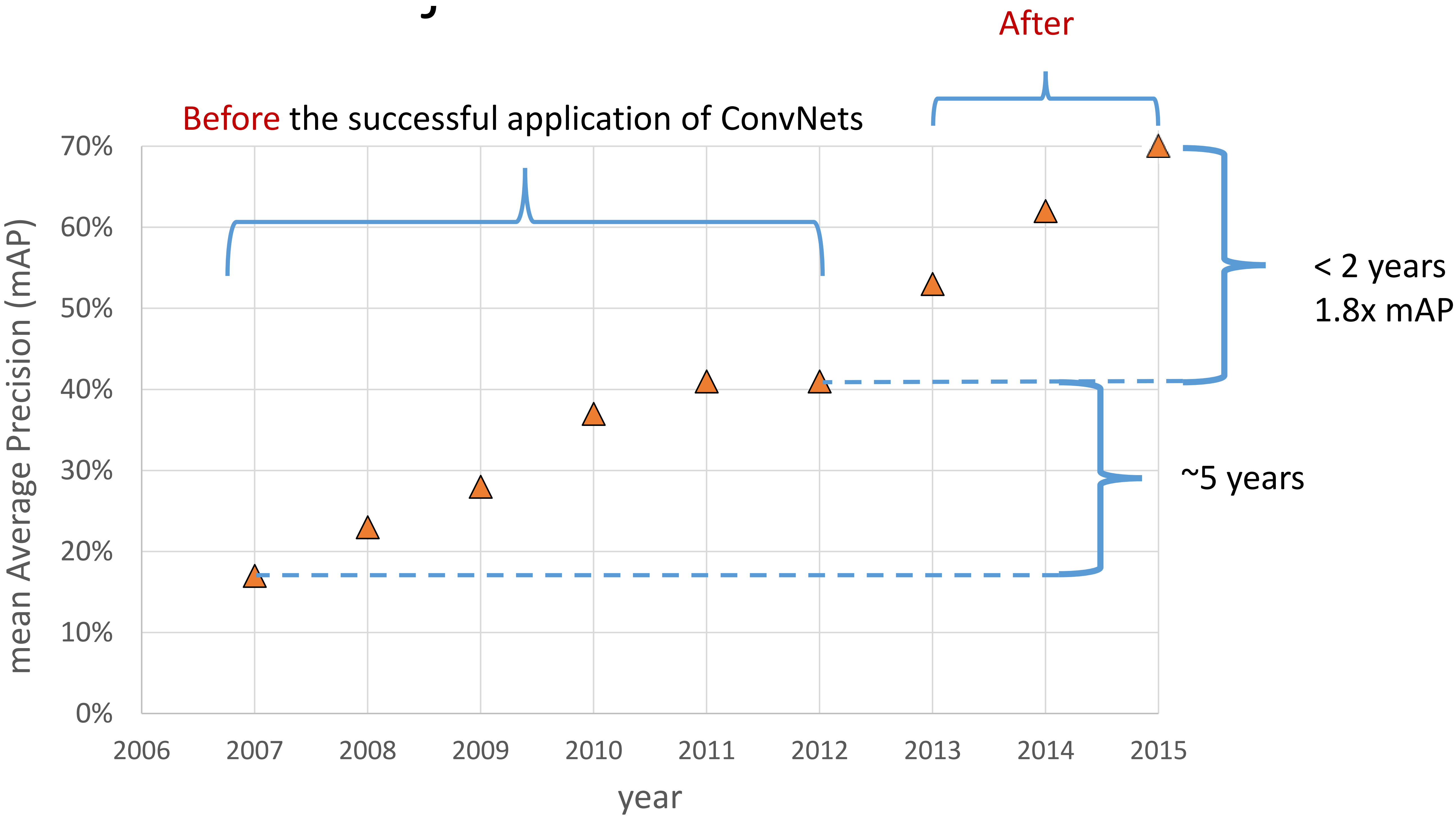
VGGnet
19 layers

GoogLeNet
22 layers

ResNet
152 layers



Pascal VOC object detection challenge



Demo

- convnet demo from Karpathy: https://cs.stanford.edu/people/karpathy/convnetjs/demo/cifar10.html

Next lecture

- gradient learning (what make it tough)
- other layers:
    - activation function,
    - batch normalization,
    - drop out,
    - loss layers